A Parallel Inertial Proximal Optimization Method

Jean-Christophe Pesquet, Nelly Pustelnik

A Parallel Inertial Proximal Optimization Method

Jean-Christophe Pesquet[1] and Nelly Pustelnik[2]

[1] Université Paris-Est
Laboratoire d'Informatique Gaspard Monge, UMR CNRS 8049
5 bd Descartes, 77454 Marne la Vallée Cedex 2, France
jean-christophe.pesquet@univ-paris-est.fr

[2] ENS Lyon
Laboratoire de Physique - CNRS UMR 5672
46 allée d'Italie, F69007 Lyon, France
nelly.pustelnik@ens-lyon.fr

Abstract

The Douglas-Rachford algorithm is a popular iterative method for finding a zero of a sum of two maximally monotone operators defined on a Hilbert space. In this paper, we propose an extension of this algorithm including inertia parameters and develop parallel versions to deal with the case of a sum of an arbitrary number of maximal operators. Based on this algorithm, parallel proximal algorithms are proposed to minimize over a linear subspace of a Hilbert space the sum of a finite number of proper, lower semicontinuous convex functions composed with linear operators. It is shown that particular cases of these methods are the simultaneous direction method of multipliers proposed by Stetzer et al., the parallel proximal algorithm developed by Combettes and Pesquet, and a parallelized version of an algorithm proposed by Attouch and Soueycatt.

Keywords: Monotone operators, convex optimization, proximal algorithms, parallel algorithms.

Mathematical Subject Classifications: 47H05, 47N10, 65K10, 68W10.

1 Introduction

The use of parallel methods for convex optimization has been an active research area for about two decades [9, 12, 34]. With the widespread use of multicore computer architectures, it can be expected that parallel optimization algorithms will play a more prominent role. Recently, a number of proximal parallel convex optimization algorithms have appeared in the literature [22]. These algorithms are especially useful for solving large-size optimization problems arising in the fields of inverse problems and imaging.

A splitting algorithm which will be subsequently designated by the *simultaneous direction method of multipliers* (SDMM) was recently proposed in [45] to solve the following problem:

$$\underset{y\in\mathcal{G}}{\text{minimize}} \quad \sum_{i=1}^{m} f_i(L_i y) \tag{1.1}$$

where $\mathcal{G}$ is a real Hilbert space, for every $i \in \{1, \ldots, m\}$, f_i is a proper lower semicontinuous convex functions from a real Hilbert space $\mathcal{H}_i$ to $]-\infty, +\infty]$, and L_i belongs to $\mathcal{B}(\mathcal{G}, \mathcal{H}_i)$, the set of bounded linear operators from $\mathcal{G}$ to $\mathcal{H}_i$. This algorithm was derived from the Douglas-Rachford algorithm [17, 19, 25, 35] by invoking a duality argument. Such a duality argument can be traced back to the work in [31] (see also [25, 48]). It was also pointed out in [45] that the proposed method generalizes to a sum of more than two functions the alternating-direction method of multipliers [30, 32]. Note that parallel alternating-direction algorithms had been previously proposed in [24] for monotropic convex programming problems and for block-separable convex optimization problems in finite dimensional Hilbert spaces. It is also worth noticing that augmented Lagrangian techniques have become increasingly popular for solving imaging problems. These approaches appeared under different names such as alternating split Bregman algorithm [29, 33, 51, 52] or split augmented Lagrangian shrinkage algorithm (SALSA) [1, 2] in the recent image processing literature.

On the other hand, another splitting method called the *parallel proximal algorithm* (PPXA) was proposed in [21] to minimize a sum of convex functions in a possibly infinite dimensional Hilbert space (see also extensions in [18] for monotone inclusion problems). Since the splitting algorithms in [21, 45] were derived from the Douglas-Rachford algorithm by working in a product space, a natural question is to know whether there exist connections between them.

Despite the fact that the work in [45] puts emphasis on duality issues whereas [21] relies on results from the theory of proximity operators, this paper shows that both algorithms are particular instances of a more general splitting algorithm. This optimization algorithm itself is a specialization of an algorithm for solving the following problem:

$$\text{find } \widetilde{y} \in E \quad \text{such that} \quad \Big(\sum_{i=1}^{m} \omega_i\, L_i^* A_i(L_i \widetilde{y})\Big) \cap E^{\perp} \neq \varnothing \tag{1.2}$$

where E is a closed linear subspace of $\mathcal{G}$, $E^{\perp}$ is its orthogonal complement and, for every $i \in \{1, \ldots, m\}$, $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ is a maximally monotone operator, L_i^* is the adjoint operator of L_i, and $\omega_i \in]0, +\infty[$. The latter algorithm for finding a zero of a sum of maximally monotone operators is derived from an extension of the Douglas-Rachford algorithm which includes inertia parameters. Note that another framework for splitting sums of maximally monotone operators was developed in [27]. Other parallel methods based on a forward-backward approach were also investigated in [5].

Convergence results concerning various inertial algorithms can be found in [3, 4, 36, 37, 38, 40, 41, 42]. Recently, a proximal alternating direction of multipliers method was introduced in [5] which can be viewed as an algorithm within this class. Applications to game theory, PDEs and control were described in [10].

Conditions for the convergence of inexact forms of the algorithms presented in this paper will be given in the following. These convergence results are valid in any finite or infinite dimensional Hilbert space. First versions of the algorithms are provided, the convergence of which requires that $\sum_{i=1}^{m} \omega_i L_i^* L_i$ is an isomorphism. This assumption is relaxed in slightly more complex variants of the algorithms. The resulting unrelaxed optimization method is shown to correspond to a parallelized version of the algorithm in [5] for equal values of the inertia parameters.

In Section 2, we introduce our notation and recall some useful properties of monotone operators. In Section 3, we propose a variant of the Douglas-Rachford algorithm in a product space, which includes inertia parameters, and study its convergence. In Section 4, we deduce from this algorithm an inertial algorithm for solving Problem (1.2). In Section 5, we consider an application of this algorithm to convex minimization problems and examine the connections of the resulting parallel convex optimization methods with some existing approaches.

2 Notation

Let $(\mathcal{H}_i, \|\cdot\|_i)_{1\le i\le m}$ be real Hilbert spaces. We define the product space $\boldsymbol{\mathcal{H}} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$ endowed with the norm

$$|||\cdot|||\colon (x_1,\ldots,x_m) \mapsto \sqrt{\sum_{i=1}^m \omega_i \|x_i\|_i^2} \tag{2.1}$$

where $(\omega_i)_{1\le i\le m} \in]0,+\infty[^m$. The associated scalar product is denoted by $\langle\langle \cdot \mid \cdot \rangle\rangle$ whereas, for every $i \in \{1,\ldots,m\}$, $\langle \cdot \mid \cdot \rangle_i$ denotes the scalar product of $\mathcal{H}_i$. Let $k \in \{0,\ldots,m\}$ and let us define the following closed subspace of $\boldsymbol{\mathcal{H}}$:

$$\boldsymbol{\mathcal{K}}_k = \{(x_1,\ldots,x_k,0,\ldots,0) \in \boldsymbol{\mathcal{H}}\} \tag{2.2}$$

(with the convention $(x_1,\ldots,x_k,0,\ldots,0) = \mathbf{0}$ if $k=0$). In the following $(x_1,\ldots,x_m) \in \mathcal{H}_1\times\cdots\times\mathcal{H}_m$ will denote the components of a generic element $\boldsymbol{x}$ of $\boldsymbol{\mathcal{H}}$.

An operator $\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ is monotone if

$$(\forall(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}}^2)(\forall(\boldsymbol{u},\boldsymbol{v}) \in \boldsymbol{A}(\boldsymbol{x})\times\boldsymbol{A}(\boldsymbol{y})) \qquad \langle\langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{u}-\boldsymbol{v}\rangle\rangle \ge 0. \tag{2.3}$$

For more details concerning the properties of monotone operators and the definitions recalled below the reader is referred to [7].

The set of zeros of $\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ is $\operatorname{zer}\boldsymbol{A} = \big\{\boldsymbol{x}\in\boldsymbol{\mathcal{H}} \;\big|\; \mathbf{0}\in\boldsymbol{A}(\boldsymbol{x})\big\}$ and its graph is $\operatorname{gra}\boldsymbol{A} = \big\{(\boldsymbol{x},\boldsymbol{u})\in\boldsymbol{\mathcal{H}}^2 \;\big|\; \boldsymbol{u}\in\boldsymbol{A}\boldsymbol{x}\big\}$. $\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ is maximally monotone if, for every $(\boldsymbol{x},\boldsymbol{u}) \in \boldsymbol{\mathcal{H}}^2$,

$$(\boldsymbol{x},\boldsymbol{u}) \in \operatorname{gra}\boldsymbol{A} \quad\Leftrightarrow\quad (\forall(\boldsymbol{y},\boldsymbol{v})\in\operatorname{gra}\boldsymbol{A})\;\; \langle\langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{u}-\boldsymbol{v}\rangle\rangle \ge 0. \tag{2.4}$$

Recall that an operator $\boldsymbol{T}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ is

- β-cocoercive with $\beta \in]0,+\infty[$ if

$$(\forall(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}}^2) \qquad \beta|||\boldsymbol{T}\boldsymbol{x}-\boldsymbol{T}\boldsymbol{y}|||^2 \le \langle\langle \boldsymbol{T}\boldsymbol{x}-\boldsymbol{T}\boldsymbol{y} \mid \boldsymbol{x}-\boldsymbol{y}\rangle\rangle; \tag{2.5}$$

- firmly nonexpansive if it is 1-cocoercive;
- $1/\beta$-Lipschitz continuous with $\beta \in]0,+\infty[$ if

$$(\forall(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}}^2) \qquad \beta|||\boldsymbol{T}\boldsymbol{x}-\boldsymbol{T}\boldsymbol{y}||| \le |||\boldsymbol{x}-\boldsymbol{y}|||; \tag{2.6}$$

- nonexpansive if it is 1-Lipschitz continuous;
- $1/\beta$-strictly contractive if it is $1/\beta$-Lipschitz continuous with $\beta \in]1,+\infty[$.

If $\boldsymbol{A}$ is maximally monotone, then its resolvent $J_{\boldsymbol{A}} = (\mathrm{Id} + \boldsymbol{A})^{-1}$ is a firmly nonexpansive operator from $\boldsymbol{\mathcal{H}}$ to $\boldsymbol{\mathcal{H}}$, and its reflection $R_{\boldsymbol{A}} = 2J_{\boldsymbol{A}} - \mathrm{Id}$ is a nonexpansive operator.

Let $\Gamma_0(\boldsymbol{\mathcal{H}})$ be the set of proper (i.e. not identically equal to $+\infty$) lower semicontinuous convex functions from $\boldsymbol{\mathcal{H}}$ to $]-\infty, +\infty]$ and let $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{H}})$. Then, its subdifferential $\partial \boldsymbol{f}$ is a maximally monotone operator and the proximity operator of $\boldsymbol{f}$ is $\mathrm{prox}_{\boldsymbol{f}} = J_{\partial \boldsymbol{f}}$ [39, 44]. The domain of $\boldsymbol{f}$ is $\mathrm{dom}\, \boldsymbol{f} = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{H}} \;\big|\; \boldsymbol{f}(\boldsymbol{x}) < +\infty\big\}$. The conjugate of $\boldsymbol{f}$ is $\boldsymbol{f}^*\colon \boldsymbol{\mathcal{H}} \to]-\infty, +\infty] : \boldsymbol{x} \mapsto \sup_{\boldsymbol{y} \in \boldsymbol{\mathcal{H}}} \langle\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle\rangle - \boldsymbol{f}(\boldsymbol{y})$.

Let $\boldsymbol{C}$ be a nonempty closed convex subset of $\boldsymbol{\mathcal{H}}$. Its indicator function $\iota_{\boldsymbol{C}} \in \Gamma_0(\boldsymbol{\mathcal{H}})$ is defined as

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad \iota_{\boldsymbol{C}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in \boldsymbol{C} \\ +\infty & \text{otherwise.} \end{cases} \tag{2.7}$$

The projection onto $\boldsymbol{C}$ is $P_{\boldsymbol{C}} = \mathrm{prox}_{\iota_{\boldsymbol{C}}}$ and the normal cone operator to $\boldsymbol{C}$ is $N_{\boldsymbol{C}}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ defined as

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad N_{\boldsymbol{C}}(\boldsymbol{x}) = \begin{cases} \big\{\boldsymbol{u} \in \boldsymbol{\mathcal{H}} \;\big|\; (\forall \boldsymbol{y} \in \boldsymbol{C}) \;\; \langle\langle \boldsymbol{y} - \boldsymbol{x} \mid \boldsymbol{u}\rangle\rangle \le 0\big\} & \text{if } \boldsymbol{x} \in \boldsymbol{C} \\ \varnothing & \text{otherwise.} \end{cases} \tag{2.8}$$

The relative interior (resp. strong relative interior) of C is denoted by $\mathrm{ri}\, C$ (resp. $\mathrm{sri}\, C$).

The weak convergence (resp. strong convergence) is denoted by $\rightharpoonup$ (resp. $\to$).

3 An extension of the Douglas-Rachford algorithm

We will first consider the following problem:

$$\text{find } \;\widetilde{\boldsymbol{x}} \in \mathrm{zer}(\boldsymbol{A} + \boldsymbol{B}) = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{H}} \;\big|\; \boldsymbol{0} \in \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{B}(\boldsymbol{x})\big\} \tag{3.1}$$

where $\boldsymbol{B}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ is a maximally monotone operator,

$$\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}\colon (x_1, \ldots, x_m) \mapsto A_1(x_1) \times \cdots \times A_m(x_m) \tag{3.2}$$

and, for every $i \in \{1, \ldots, m\}$, $A_i\colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ is a maximally monotone operator.

We first state a quasi-Fejérian property which will be useful in the following.

Lemma 3.1 *Let $\boldsymbol{S}$ be a nonempty set of $\boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{K}}_k$. Let $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ be a sequence in $\boldsymbol{\mathcal{H}}$ and let $(\boldsymbol{p}_n)_{n\in\mathbb{N}} = \big((p_{1,n}, \ldots, p_{k,n}, 0, \ldots, 0)\big)_{n\in\mathbb{N}}$ be a sequence in $\boldsymbol{\mathcal{K}}_k$ such that*

$$\begin{aligned} &(\forall (\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}}) \in \boldsymbol{S})(\forall n \in \mathbb{N}) \\ &\quad \Big(|||\boldsymbol{t}_{n+1} - \widetilde{\boldsymbol{t}}|||^2 + \sum_{i=1}^{k} \mu_{i,n+1} \|p_{i,n+1} - \widetilde{p}_i\|_i^2\Big)^{1/2} \le \Big(|||\boldsymbol{t}_n - \widetilde{\boldsymbol{t}}|||^2 + \sum_{i=1}^{k} \mu_{i,n} \|p_{i,n} - \widetilde{p}_i\|_i^2\Big)^{1/2} + \Delta_n \end{aligned} \tag{3.3}$$

where, for every $i \in \{1, \ldots, k\}$, $(\mu_{i,n})_{n\in\mathbb{N}}$ is a sequence of nonnegative reals converging to a positive limit, and $(\Delta_n)_{n\in\mathbb{N}}$ is a summable sequence of nonnegative reals. Then,

(i) *$(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ are bounded sequences.*

(ii) *For every* $(\widetilde{\boldsymbol{t}},\widetilde{\boldsymbol{p}})\in\boldsymbol{S}$, $(|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\sum_{i=1}^k\mu_{i,n}\|p_{i,n}-\widetilde{p}_i\|_i^2)_{n\in\mathbb{N}}$ *converges.*

(iii) *If every weak sequential cluster point of* $(\boldsymbol{t}_n,\boldsymbol{p}_n)_{n\in\mathbb{N}}$ *belongs to* $\boldsymbol{S}$*, then* $(\boldsymbol{t}_n,\boldsymbol{p}_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\boldsymbol{S}$*.*

Proof.

(i) Let $(\widetilde{\boldsymbol{t}},\widetilde{\boldsymbol{p}})\in\boldsymbol{S}$. According to [16, Lemma 3.1(i)], there exists $\rho\in\,]0,+\infty[$ such that

$$(\forall n\in\mathbb{N})\quad\Big(|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\sum_{i=1}^k\mu_{i,n}\|p_{i,n}-\widetilde{p}_i\|_i^2\Big)^{1/2}\le\rho.\tag{3.4}$$

Since, for every $i\in\{1,\ldots,k\}$, $(\mu_{i,n})_{n\in\mathbb{N}}$ is a sequence of nonnegative reals converging to a positive limit,

$$\exists\underline{\mu}\in\,]0,+\infty[\,,\exists n_0\in\mathbb{N}\quad(\forall i\in\{1,\ldots,k\})(\forall n\ge n_0)\quad\mu_{i,n}\ge\underline{\mu}\,\omega_i\tag{3.5}$$

It can be deduced that

$$(\forall n\ge n_0)\quad\Big(|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\underline{\mu}|||\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}|||^2\Big)^{1/2}\le\rho.\tag{3.6}$$

This implies that

$$(\forall n\ge n_0)\qquad\begin{aligned}|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||\ &\le\rho\\|||\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}|||\ &\le\underline{\mu}^{-1/2}\rho\end{aligned}\tag{3.7}$$

which shows the boundedness of the sequences $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$.

(ii) This fact follows from [16, Lemma 3.1 (ii)].

(iii) Since $(\boldsymbol{t}_n,\boldsymbol{p}_n)_{n\in\mathbb{N}}$ is bounded, it is enough to prove that this sequence cannot have two distinct weak sequential cluster points. Let $(\widetilde{\boldsymbol{t}}_1,\widetilde{\boldsymbol{p}}_1)\in\boldsymbol{S}$ and $(\widetilde{\boldsymbol{t}}_2,\widetilde{\boldsymbol{p}}_2)\in\boldsymbol{S}$ be two such cluster points. We have

$$(\forall n\in\mathbb{N})\qquad 2\langle\langle\boldsymbol{t}_n\mid\widetilde{\boldsymbol{t}}_1-\widetilde{\boldsymbol{t}}_2\rangle\rangle=|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}_2|||^2-|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}_1|||^2+|||\widetilde{\boldsymbol{t}}_1|||^2-|||\widetilde{\boldsymbol{t}}_2|||^2$$

and

$$(\forall i\in\{1,\ldots,k\})(\forall n\in\mathbb{N})\quad 2\langle p_{i,n}\mid\widetilde{p}_{i,1}-\widetilde{p}_{i,2}\rangle_i=\|p_{i,n}-\widetilde{p}_{i,2}\|_i^2-\|p_{i,n}-\widetilde{p}_{i,1}\|_i^2+\|\widetilde{p}_{i,1}\|_i^2-\|\widetilde{p}_{i,2}\|_i^2.\tag{3.8}$$

Then, it can be deduced from (ii) and the fact that, for every $i\in\{1,\ldots,k\}$, $(\mu_{i,n})_{n\in\mathbb{N}}$ is convergent that $(\langle\langle\boldsymbol{t}_n\mid\widetilde{\boldsymbol{t}}_1-\widetilde{\boldsymbol{t}}_2\rangle\rangle+\sum_{i=1}^k\mu_{i,n}\langle p_{i,n}\mid\widetilde{p}_{i,1}-\widetilde{p}_{i,2}\rangle_i)_{n\in\mathbb{N}}$ converges to some limit ζ. In addition, there exist some subsequences $(\boldsymbol{t}_{n_\ell},\boldsymbol{p}_{n_\ell})_{\ell\in\mathbb{N}}$ and $(\boldsymbol{t}_{n_{\ell'}},\boldsymbol{p}_{n_{\ell'}})_{\ell'\in\mathbb{N}}$ converging weakly to $(\widetilde{\boldsymbol{t}}_1,\widetilde{\boldsymbol{p}}_1)$ and $(\widetilde{\boldsymbol{t}}_2,\widetilde{\boldsymbol{p}}_2)$, respectively. Passing to the limit, we have thus

$$\begin{aligned}&\langle\langle\widetilde{\boldsymbol{t}}_1\mid\widetilde{\boldsymbol{t}}_1-\widetilde{\boldsymbol{t}}_2\rangle\rangle+\sum_{i=1}^k\mu_{i,\infty}\langle\widetilde{p}_{i,1}\mid\widetilde{p}_{i,1}-\widetilde{p}_{i,2}\rangle_i=\zeta=\langle\langle\widetilde{\boldsymbol{t}}_2\mid\widetilde{\boldsymbol{t}}_1-\widetilde{\boldsymbol{t}}_2\rangle\rangle+\sum_{i=1}^k\mu_{i,\infty}\langle\widetilde{p}_{i,2}\mid\widetilde{p}_{i,1}-\widetilde{p}_{i,2}\rangle_i\\&\Leftrightarrow|||\widetilde{\boldsymbol{t}}_1-\widetilde{\boldsymbol{t}}_2|||^2+\sum_{i=1}^k\mu_{i,\infty}\|\widetilde{p}_{i,1}-\widetilde{p}_{i,2}\|_i^2=0\end{aligned}\tag{3.9}$$

where, for every $i\in\{1,\ldots,k\}$, $\mu_{i,\infty}=\lim_{n\to+\infty}\mu_{i,n}>0$. Consequently, $\widetilde{\boldsymbol{t}}_1=\widetilde{\boldsymbol{t}}_2$ and $\widetilde{\boldsymbol{p}}_1=\widetilde{\boldsymbol{p}}_2$.

□

We will also need the following property concerning nonexpansive operators.

Lemma 3.2 *Let* $(\gamma, \lambda) \in]0, +\infty[^2$ *and let* $(\varepsilon_1, \ldots, \varepsilon_m) \in [0, 1[^m$ *be such that*

$$(\forall i \in \{1, \ldots, m\}) \qquad (\varepsilon_i > 0 \Leftrightarrow i \leq k). \tag{3.10}$$

Let $\boldsymbol{T}$ *and* $\boldsymbol{U}_\lambda$ *be the operators from* $\boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{K}}_k$ *to* $\boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{K}}_k$ *defined as*

$$(\forall \boldsymbol{t} \in \boldsymbol{\mathcal{H}})(\forall \boldsymbol{p} \in \boldsymbol{\mathcal{K}}_k) \quad \boldsymbol{T}(\boldsymbol{t}, \boldsymbol{p}) = (\boldsymbol{z}, P_{\boldsymbol{\mathcal{K}}_k} \boldsymbol{p}') \tag{3.11}$$

$$\boldsymbol{U}_\lambda(\boldsymbol{t}, \boldsymbol{p}) = \Big((1 - \frac{\lambda}{2})\boldsymbol{t} + \frac{\lambda}{2}\boldsymbol{z}, P_{\boldsymbol{\mathcal{K}}_k} \boldsymbol{p}'\Big) \tag{3.12}$$

where

$$\boldsymbol{p}' = \Big(J_{\gamma(1-\varepsilon_1)A_1}\big((1 - \varepsilon_1)t_1 + \varepsilon_1 p_1\big), \ldots, J_{\gamma(1-\varepsilon_m)A_m}\big((1 - \varepsilon_m)t_m + \varepsilon_m p_m\big)\Big) \tag{3.13}$$

$$\boldsymbol{z} = R_{\gamma \boldsymbol{B}}\big(2\boldsymbol{p}' - \boldsymbol{t}\big). \tag{3.14}$$

We have the following properties:

(i) $\operatorname{Fix} \boldsymbol{T} = \operatorname{Fix} \boldsymbol{U}_\lambda$, *and*

$$\begin{aligned} &(\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}}) \in \operatorname{Fix} \boldsymbol{T} \\ \Leftrightarrow \;& J_{\gamma \boldsymbol{A}} \widetilde{\boldsymbol{t}} \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \text{ and } (\forall i \in \{1, \ldots, m\}) \; \widetilde{p}_i = \begin{cases} J_{\gamma A_i} \widetilde{t}_i & \text{if } i \leq k \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{3.15}$$

(ii) $\boldsymbol{T}$ *is nonexpansive operator in* $\boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{K}}_k$ *endowed with the weighted norm*

$$\nu \colon (\boldsymbol{z}, \boldsymbol{p}) \mapsto \Big(|||\boldsymbol{z}|||^2 + 2 \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1 - \varepsilon_i} \|p_i\|_i^2 \Big)^{1/2}. \tag{3.16}$$

(iii) *For every* $(\boldsymbol{t}, \boldsymbol{s}) \in \boldsymbol{\mathcal{H}}^2$ *and* $(\boldsymbol{p}, \boldsymbol{q}) \in \boldsymbol{\mathcal{K}}_k^2$, *let* $(\boldsymbol{z}, \overline{\boldsymbol{p}}) = \boldsymbol{T}(\boldsymbol{t}, \boldsymbol{p})$ *and* $(\boldsymbol{u}, \overline{\boldsymbol{q}}) = \boldsymbol{T}(\boldsymbol{s}, \boldsymbol{q})$.

Then,

$$\begin{aligned} |||(1 - \frac{\lambda}{2})(\boldsymbol{t} - \boldsymbol{s}) + \frac{\lambda}{2}(\boldsymbol{z} - \boldsymbol{u})|||^2 + \lambda \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1 - \varepsilon_i} \|\overline{p}_i - \overline{q}_i\|_i^2 \leq |||\boldsymbol{t} - \boldsymbol{s}|||^2 + \lambda \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1 - \varepsilon_i} \|p_i - q_i\|_i^2 \\ - \lambda \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1 - \varepsilon_i} \|\overline{p}_i - \overline{q}_i - p_i + q_i\|_i^2 - \frac{1}{4}(2 - \lambda)\lambda |||\boldsymbol{t} - \boldsymbol{s} - \boldsymbol{z} + \boldsymbol{u}|||^2. \end{aligned} \tag{3.17}$$

Proof.

(i) It can be first noticed that $\operatorname{Fix} \boldsymbol{T} = \operatorname{Fix} \boldsymbol{U}_\lambda$. In addition, $(\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}})$ is a fixed point of $\boldsymbol{T}$ if and only if

$$\begin{cases} (\forall i \in \{1, \ldots, m\}) \quad \widetilde{p}_i = \begin{cases} \widetilde{p}'_i & \text{if } i \leq k \\ 0 & \text{otherwise} \end{cases} \\ \widetilde{\boldsymbol{t}} = R_{\gamma \boldsymbol{B}}\big(2\widetilde{\boldsymbol{p}}' - \widetilde{\boldsymbol{t}}\big) \end{cases} \tag{3.18}$$

where

$$(\forall i \in \{1,\dots,m\}) \qquad \widetilde{p}'_i = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)\widetilde{t}_i + \varepsilon_i \widetilde{p}_i\big). \tag{3.19}$$

The latter relation yields, for every $i \in \{1,\dots,k\}$,

$$\begin{aligned}
&(1-\varepsilon_i)\widetilde{t}_i + \varepsilon_i \widetilde{p}'_i - \widetilde{p}'_i \in \gamma(1-\varepsilon_i)A_i(\widetilde{p}'_i)\\
\Leftrightarrow\quad & \widetilde{t}_i - \widetilde{p}'_i \in \gamma A_i(\widetilde{p}'_i)\\
\Leftrightarrow\quad & \widetilde{p}'_i = J_{\gamma A_i}\widetilde{t}_i.
\end{aligned} \tag{3.20}$$

So, by using (3.10), (3.19) is equivalent to

$$\widetilde{\boldsymbol{p}}' = J_{\gamma \boldsymbol{A}}\widetilde{\boldsymbol{t}}. \tag{3.21}$$

and $(\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}}) \in \operatorname{Fix}\boldsymbol{T}$ if and only if

$$(\forall i \in \{1,\dots,m\}) \quad \widetilde{p}_i = \begin{cases} J_{\gamma A_i}\widetilde{t}_i & \text{if } i \le k\\ 0 & \text{otherwise}\end{cases} \tag{3.22}$$

$$\widetilde{\boldsymbol{t}} = R_{\gamma \boldsymbol{B}}\big(2\, J_{\gamma \boldsymbol{A}}\widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{t}}\big). \tag{3.23}$$

On the other hand, $\widetilde{\boldsymbol{t}}$ satisfies (3.23) if and only if it is a fixed point of the standard Douglas-Rachford iteration, that is if and only if $J_{\gamma \boldsymbol{A}}\widetilde{\boldsymbol{t}}$ is a zero of $\boldsymbol{A} + \boldsymbol{B}$ [17, 35].

(ii) For every $(\boldsymbol{t}, \boldsymbol{s}) \in \boldsymbol{\mathcal{H}}^2$ and $(\boldsymbol{p}, \boldsymbol{q}) \in \boldsymbol{\mathcal{K}}_k^2$, let $(\boldsymbol{z}, \overline{\boldsymbol{p}}) = \boldsymbol{T}(\boldsymbol{t}, \boldsymbol{p})$ and $(\boldsymbol{u}, \overline{\boldsymbol{q}}) = \boldsymbol{T}(\boldsymbol{s}, \boldsymbol{q})$. Let $\boldsymbol{p}'$ be given by (3.13) and $\boldsymbol{q}'$ be similarly defined as

$$\boldsymbol{q}' = \big(J_{\gamma(1-\varepsilon_1)A_1}\big((1-\varepsilon_1)s_1 + \varepsilon_1 q_1\big), \dots, J_{\gamma(1-\varepsilon_m)A_m}\big((1-\varepsilon_m)s_m + \varepsilon_m q_m\big)\big). \tag{3.24}$$

For every $i \in \{1,\dots,m\}$, as $J_{\gamma(1-\varepsilon_i)A_i}$ is firmly nonexpansive, we have

$$\begin{aligned}
&\|p'_i - q'_i\|_i^2 \le \langle p'_i - q'_i \mid (1-\varepsilon_i)(t_i - s_i) + \varepsilon_i(p_i - q_i)\rangle_i\\
\Leftrightarrow\quad & \|p'_i - q'_i\|_i^2 \le (1-\varepsilon_i)\langle p'_i - q'_i \mid t_i - s_i\rangle_i + \frac{\varepsilon_i}{2}(\|p'_i - q'_i\|_i^2 + \|p_i - q_i\|_i^2\\
&\qquad - \|p'_i - q'_i - p_i + q_i\|_i^2)\\
\Leftrightarrow\quad & -(1-\varepsilon_i)\langle p'_i - q'_i \mid t_i - s_i\rangle_i \le -\Big(1 - \frac{\varepsilon_i}{2}\Big)\|p'_i - q'_i\|_i^2 + \frac{\varepsilon_i}{2}\|p_i - q_i\|_i^2 - \frac{\varepsilon_i}{2}\|p'_i - q'_i - p_i + q_i\|_i^2.
\end{aligned} \tag{3.25}$$

On the other hand, we deduce from the nonexpansivity of $R_{\gamma \boldsymbol{B}}$ that

$$\begin{aligned}
&|||\boldsymbol{z} - \boldsymbol{u}|||^2 \le |||2(\boldsymbol{p}' - \boldsymbol{q}') - \boldsymbol{t} + \boldsymbol{s}|||^2\\
\Leftrightarrow\quad & |||\boldsymbol{z} - \boldsymbol{u}|||^2 \le 4|||\boldsymbol{p}' - \boldsymbol{q}'|||^2 + |||\boldsymbol{t} - \boldsymbol{s}|||^2 - 4\langle\langle \boldsymbol{p}' - \boldsymbol{q}' \mid \boldsymbol{t} - \boldsymbol{s}\rangle\rangle.
\end{aligned} \tag{3.26}$$

By combining this inequality with (3.25), we get

$$\begin{aligned}
&|||\boldsymbol{z} - \boldsymbol{u}|||^2 + 2\sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i}\|p'_i - q'_i\|_i^2\\
&\qquad \le |||\boldsymbol{t} - \boldsymbol{s}|||^2 + 2\sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i}\|p_i - q_i\|_i^2 - 2\sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i}\ \|p'_i - q'_i - p_i + q_i\|_i^2.
\end{aligned} \tag{3.27}$$

This leads to

$$|||\boldsymbol{z}-\boldsymbol{u}|||^2+2\sum_{i=1}^{k}\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_i'-q_i'\|_i^2\le|||\boldsymbol{t}-\boldsymbol{s}|||^2+2\sum_{i=1}^{k}\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_i-q_i\|_i^2. \tag{3.28}$$

Since $\overline{\boldsymbol{p}}=(p_1',\ldots,p_k',0,\ldots,0)$ and $\overline{\boldsymbol{q}}=(q_1',\ldots,q_k',0,\ldots,0)$, the above inequality implies that $\boldsymbol{T}$ is a nonexpansive operator, as stated above.

(iii) In addition, we have

$$\begin{aligned}
&|||(1-\frac{\lambda}{2})(\boldsymbol{t}-\boldsymbol{s})+\frac{\lambda}{2}(\boldsymbol{z}-\boldsymbol{u})|||^2\\
&=(1-\frac{\lambda}{2})^2|||\boldsymbol{t}-\boldsymbol{s}|||^2+\frac{\lambda^2}{4}|||\boldsymbol{z}-\boldsymbol{u}|||^2+\frac{1}{2}(2-\lambda)\lambda\langle\langle\boldsymbol{t}-\boldsymbol{s}\mid\boldsymbol{z}-\boldsymbol{u}\rangle\rangle\\
&=(1-\frac{\lambda}{2})|||\boldsymbol{t}-\boldsymbol{s}|||^2+\frac{\lambda}{2}|||\boldsymbol{z}-\boldsymbol{u}|||^2-\frac{1}{4}(2-\lambda)\lambda|||\boldsymbol{t}-\boldsymbol{s}-\boldsymbol{z}+\boldsymbol{u}|||^2
\end{aligned} \tag{3.29}$$

which, combined with (3.27), yields (3.17).

□

We are now able to provide the main result of this section, which concerns the convergence of a generalization of the Douglas-Rachford algorithm.

Proposition 3.3 *Let $\gamma\in\,]0,+\infty[$, $(\varepsilon_1,\ldots,\varepsilon_m)\in[0,1[^m$ and $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of reals. Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ be sequences in $\boldsymbol{\mathcal{H}}$, and let $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{p}_n)_{n\ge-1}$ be sequences generated by the following routine.*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\; \boldsymbol{t}_0\in\boldsymbol{\mathcal{H}},\boldsymbol{p}_{-1}\in\boldsymbol{\mathcal{H}}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
\text{For } i=1,\ldots,m\\
\quad\left\lfloor\; p_{i,n}=J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n}+\varepsilon_i p_{i,n-1}\big)+a_{i,n}\right.\\
\boldsymbol{t}_{n+1}=\boldsymbol{t}_n+\lambda_n\Big(J_{\gamma\boldsymbol{B}}\big(2\boldsymbol{p}_n-\boldsymbol{t}_n\big)+\boldsymbol{b}_n-\boldsymbol{p}_n\Big).
\end{array}\right.
\end{array} \tag{3.30}$$

Suppose that the following hold.

(i) $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})\neq\varnothing$

(ii) *There exists $\underline{\lambda}\in]0,2[$ such that $(\forall n\in\mathbb{N})$ $\underline{\lambda}\le\lambda_{n+1}\le\lambda_n<2$.*

(iii) $\sum_{n\in\mathbb{N}}|||\boldsymbol{a}_n|||+|||\boldsymbol{b}_n|||<+\infty$.

Then, $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ converges weakly to $\widetilde{\boldsymbol{t}}$ and $J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

Proof. Without loss of generality, it can be assumed that (3.10) is satisfied (up to some re-indexing). For every $n\in\mathbb{N}$, let

$$\boldsymbol{p}_n'=\big(J_{\gamma(1-\varepsilon_1)A_1}\big((1-\varepsilon_1)t_{1,n}+\varepsilon_1p_{1,n-1}\big),\ldots,J_{\gamma(1-\varepsilon_m)A_m}\big((1-\varepsilon_m)t_{m,n}+\varepsilon_mp_{m,n-1}\big)\big) \tag{3.31}$$

$$\boldsymbol{z}_n=R_{\gamma\boldsymbol{B}}\big(2\boldsymbol{p}_n'-\boldsymbol{t}_n\big) \tag{3.32}$$

$$\boldsymbol{t}_{n+1}'=(1-\frac{\lambda_n}{2})\boldsymbol{t}_n+\frac{\lambda_n}{2}\boldsymbol{z}_n. \tag{3.33}$$

We have then

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} (\boldsymbol{z}_n, P_{\boldsymbol{\mathcal{K}_k}} \boldsymbol{p}'_n) = \boldsymbol{T}(\boldsymbol{t}_n, P_{\boldsymbol{\mathcal{K}_k}} \boldsymbol{p}_{n-1}) \\ (\boldsymbol{t}'_{n+1}, P_{\boldsymbol{\mathcal{K}_k}} \boldsymbol{p}'_n) = \boldsymbol{U}_{\lambda_n}(\boldsymbol{t}_n, P_{\boldsymbol{\mathcal{K}_k}} \boldsymbol{p}_{n-1}) \end{cases} \tag{3.34}$$

where $\boldsymbol{T}$ and $\boldsymbol{U}_{\lambda_n}$ are defined by (3.11)-(3.14). According to Lemma 3.2(i) and Assumption (i), $\operatorname{Fix} \boldsymbol{U}_{\lambda_n} = \operatorname{Fix} \boldsymbol{T} \neq \varnothing$. Let $(\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}})$ be an arbitrary element of $\operatorname{Fix} \boldsymbol{T}$ and, for every $n \in \mathbb{N}$, define $\widehat{\boldsymbol{t}}_n = \boldsymbol{t}_n - \widetilde{\boldsymbol{t}}$, $\widehat{\boldsymbol{p}}_{n-1} = \boldsymbol{p}_{n-1} - \widetilde{\boldsymbol{p}}$, $\widehat{\boldsymbol{t}}'_{n+1} = \boldsymbol{t}'_{n+1} - \widetilde{\boldsymbol{t}}$, and $\widehat{\boldsymbol{p}}'_n = \boldsymbol{p}'_n - \widetilde{\boldsymbol{p}}$. By applying Lemma 3.2(iii), we get

$$\begin{aligned} &|||\widehat{\boldsymbol{t}}'_{n+1}|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}'_{i,n}\|_i^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|p'_{i,n} - p_{i,n-1}\|_i^2 + \frac{1}{4}(2-\lambda_n)\lambda_n |||\boldsymbol{t}_n - \boldsymbol{z}_n|||^2 \\ &\le |||\widehat{\boldsymbol{t}}_n|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}_{i,n-1}\|_i^2. \end{aligned} \tag{3.35}$$

From the triangular inequality, we have, for every $n \in \mathbb{N}$,

$$\begin{aligned} &\Big(|||\widehat{\boldsymbol{t}}_{n+1}|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}_{i,n}\|_i^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|p_{i,n} - p_{i,n-1}\|_i^2 + \frac{1}{4}(2-\lambda_n)\lambda_n |||\boldsymbol{t}_n - \boldsymbol{z}_n|||^2\Big)^{1/2} \\ &\le \Big(|||\widehat{\boldsymbol{t}}'_{n+1}|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}'_{i,n}\|_i^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|p'_{i,n} - p_{i,n-1}\|_i^2 + \frac{1}{4}(2-\lambda_n)\lambda_n |||\boldsymbol{t}_n - \boldsymbol{z}_n|||^2\Big)^{1/2} \\ &\quad + |||\boldsymbol{t}'_{n+1} - \boldsymbol{t}_{n+1}||| + (2\lambda_n)^{1/2} \Big(\sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|p'_{i,n} - p_{i,n}\|_i^2\Big)^{1/2}. \end{aligned} \tag{3.36}$$

By straightforward calculations, the iterations of Algorithm (3.30) can be expressed as

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} \boldsymbol{p}_n = \boldsymbol{p}'_n + \boldsymbol{a}_n \\ \boldsymbol{t}_{n+1} = \boldsymbol{t}'_{n+1} + \frac{\lambda_n}{2}\big(R_{\gamma \boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n) - \boldsymbol{z}_n + 2\boldsymbol{b}_n\big). \end{cases} \tag{3.37}$$

From the nonexpansivity of the $R_{\gamma \boldsymbol{B}}$ operator, it can be deduced that

$$(\forall n \in \mathbb{N}) \qquad |||\boldsymbol{t}_{n+1} - \boldsymbol{t}'_{n+1}||| \le \frac{\lambda_n}{2} |||R_{\gamma \boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n) - \boldsymbol{z}_n||| + \lambda_n |||\boldsymbol{b}_n||| \tag{3.38}$$

$$|||R_{\gamma \boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n) - \boldsymbol{z}_n||| \le 2|||\boldsymbol{p}_n - \boldsymbol{p}'_n||| = 2|||\boldsymbol{a}_n||| \tag{3.39}$$

and, by using (3.35), (3.36) becomes

$$\begin{aligned} &\Big(|||\widehat{\boldsymbol{t}}_{n+1}|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}_{i,n}\|_i^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|p_{i,n} - p_{i,n-1}\|_i^2 + \frac{1}{4}(2-\lambda_n)\lambda_n |||\boldsymbol{t}_n - \boldsymbol{z}_n|||^2\Big)^{1/2} \\ &\le \Big(|||\widehat{\boldsymbol{t}}_n|||^2 + \lambda_n \sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|\widehat{p}_{i,n-1}\|_i^2\Big)^{1/2} + \Delta_n \end{aligned} \tag{3.40}$$

where

$$\Delta_n = \lambda_n(|||\boldsymbol{a}_n||| + |||\boldsymbol{b}_n|||) + (2\lambda_n)^{1/2} \Big(\sum_{i=1}^k \frac{\omega_i \varepsilon_i}{1-\varepsilon_i} \|a_{i,n}\|_i^2\Big)^{1/2}. \tag{3.41}$$

Since, due to Assumption (ii), $(\lambda_n)_{n\in\mathbb{N}}$ is decreasing, (3.40) leads to

$$\Big(|||\widehat{\boldsymbol{t}}_{n+1}|||^2+\lambda_n\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|\widehat{p}_{i,n}\|_i^2\Big)^{1/2}\le\Big(|||\widehat{\boldsymbol{t}}_n|||^2+\lambda_{n-1}\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|\widehat{p}_{i,n-1}\|_i^2\Big)^{1/2}+\Delta_n. \quad (3.42)$$

Since Assumption (iii) holds and $(\lambda_n)_{n\in\mathbb{N}}$ is bounded, the above expression of $(\Delta_n)_{n\in\mathbb{N}}$ shows that it is a nonnegative summable sequence. Then, as $(\lambda_n)_{n\in\mathbb{N}}$ converges to a positive limit, Lemma 3.1(ii) allows us to claim that $(|||\widehat{\boldsymbol{t}}_n|||^2+\lambda_{n-1}\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|\widehat{p}_{i,n-1}\|_i^2)_{n\ge1}$ is a convergent sequence. In turn, as $\lim_{n\to+\infty}\Delta_n=0$, we deduce from (3.40) and (3.42) that

$$\begin{aligned}\underline{\lambda}\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n}-p_{i,n-1}\|_i^2+\frac14(2-\lambda_0)\underline{\lambda}|||\boldsymbol{t}_n-\boldsymbol{z}_n|||^2&\\ \le\lambda_n\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n}-p_{i,n-1}\|_i^2+\frac14(2-\lambda_n)\lambda_n|||\boldsymbol{t}_n-\boldsymbol{z}_n|||^2\to0,& \quad (3.43)\end{aligned}$$

By using now (3.37), we have

$$\begin{aligned}&\Big(\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p'_{i,n}-p_{i,n-1}\|_i^2+\frac14(2-\lambda_0)\,|||\boldsymbol{t}_n-\boldsymbol{z}_n|||^2\Big)^{1/2}\\ &\quad\le\Big(\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n}-p_{i,n-1}\|_i^2+\frac14(2-\lambda_0)|||\boldsymbol{t}_n-\boldsymbol{z}_n|||^2\Big)^{1/2}+\Big(\frac{\overline{\varepsilon}}{1-\overline{\varepsilon}}\Big)^{1/2}|||\boldsymbol{a}_n|||\to0 \quad (3.44)\end{aligned}$$

where $\overline{\varepsilon}=\max_{1\le i\le k}\varepsilon_i$. This entails that

$$\nu\big(\boldsymbol{T}(\boldsymbol{t}_n,P_{\boldsymbol{\mathcal{K}}_k}\boldsymbol{p}_{n-1})-(\boldsymbol{t}_n,P_{\boldsymbol{\mathcal{K}}_k}\boldsymbol{p}_{n-1})\big)\to0 \quad (3.45)$$

where ν is the norm given by (3.16). Appealing to Lemma 3.2(ii), if $(\boldsymbol{t}_{n_\ell},P_{\boldsymbol{\mathcal{K}}_k}\boldsymbol{p}_{n_\ell-1})_{\ell\in\mathbb{N}}$ is a weakly converging subsequence of $(\boldsymbol{t}_n,P_{\boldsymbol{\mathcal{K}}_k}\boldsymbol{p}_{n-1})_{n\in\mathbb{N}}$, we deduce from the demiclosedness principle [11] that its limit belongs to $\operatorname{Fix}\boldsymbol{T}$. As a result of Lemma 3.1(iii), we conclude that

$$(\boldsymbol{t}_n,P_{\boldsymbol{\mathcal{K}}_k}\boldsymbol{p}_{n-1})\rightharpoonup(\widetilde{\boldsymbol{t}},\widetilde{\boldsymbol{p}})\in\operatorname{Fix}\boldsymbol{T}. \quad (3.46)$$

By using Lemma 3.2(i), we have thus $J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. □

Remark 3.4

(i) In Algorithm (3.30), $(\lambda_n)_{n\in\mathbb{N}}$ and $(\varepsilon_1,\ldots,\varepsilon_m)$ correspond to relaxation and inertia parameters, respectively. The sequences $(a_{i,n})_{n\in\mathbb{N}}$ with $i\in\{1,\ldots,m\}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ model possible errors in the computation of $J_{\gamma(1-\varepsilon_i)A_i}$ and $J_{\gamma\boldsymbol{B}}$.

(ii) When $(\forall i\in\{1,\ldots,m\})\ \varepsilon_i=0$, weaker conditions than (ii) and (iii) are known [17] to be sufficient to prove the convergence result, namely

(a) $\sum_{n\in\mathbb{N}}\lambda_n(2-\lambda_n)=+\infty$

(b) $\sum_{n\in\mathbb{N}}\lambda_n(|||\boldsymbol{a}_n|||+|||\boldsymbol{b}_n|||)<+\infty$.

Stronger convergence results can be deduced from the previous proposition:

Proposition 3.5 *Suppose that the assumptions of Proposition 3.3 hold. Let* $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ *and* $(\boldsymbol{p}_n)_{n\geq -1}$ *be sequences generated by Algorithm* (3.30)*. Then,* $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ *and* $(\boldsymbol{y}_n)_{n\in\mathbb{N}} = \big(J_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n)\big)_{n\in\mathbb{N}}$ *both converge weakly to* $\widetilde{\boldsymbol{y}} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$*.*

Proof. We proceed similarly to the proof in [7, Theorem 25.6] for the convergence of the standard Douglas-Rachford algorithm, which is inspired from [26, 49].
By using the same notation as in the proof of Proposition 3.3, it can be deduced from (3.46) that

$$(\forall i \in \{1,\dots,m\}) \qquad (1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1} \rightharpoonup (1-\varepsilon_i)\widetilde{t}_i + \varepsilon_i \widetilde{p}_i \tag{3.47}$$

(since $i > k \Rightarrow \varepsilon_i = 0$), which implies that $\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big)_{n\in\mathbb{N}}$ is a bounded sequence. For every $i \in \{1,\dots,m\}$, as a consequence of the nonexpansivity of $J_{\gamma(1-\varepsilon_i)A_i}$, $(\boldsymbol{p}'_n)_{n\in\mathbb{N}}$ as defined by (3.31) is bounded. Hence, there exists a subsequence $(\boldsymbol{p}'_{n_\ell})_{\ell\in\mathbb{N}}$ of $(\boldsymbol{p}'_n)_{n\in\mathbb{N}}$ and $\widetilde{\boldsymbol{y}} \in \boldsymbol{\mathcal{H}}$ such that

$$\boldsymbol{p}'_{n_\ell} \rightharpoonup \widetilde{\boldsymbol{y}}. \tag{3.48}$$

Due to (3.37) and Assumption (iii) of Proposition 3.3, we have also

$$\boldsymbol{p}_{n_\ell} \rightharpoonup \widetilde{\boldsymbol{y}} \tag{3.49}$$

and, by using (3.46),

$$(\forall i \in \{1,\dots,k\}) \qquad \widetilde{y}_i = \widetilde{p}_i. \tag{3.50}$$

Furthermore, we have

$$\begin{aligned}(\forall n \in \mathbb{N}) \qquad |||\boldsymbol{p}_n - \boldsymbol{y}_n||| &= \frac{1}{2}|||\boldsymbol{t}_n - R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n)||| \\ &\leq \frac{1}{2}\big(|||\boldsymbol{t}_n - \boldsymbol{z}_n||| + |||\boldsymbol{z}_n - R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{t}_n)|||\big) \\ &\leq \frac{1}{2}|||\boldsymbol{t}_n - \boldsymbol{z}_n||| + |||\boldsymbol{a}_n||| \end{aligned} \tag{3.51}$$

where (3.39) has been used to get the last inequality. In addition, according to (3.45), $|||\boldsymbol{t}_n - \boldsymbol{z}_n||| \to 0$ and we know that $\boldsymbol{a}_n \to \boldsymbol{0}$. Hence, (3.51) yields

$$\boldsymbol{p}_n - \boldsymbol{y}_n \to \boldsymbol{0} \tag{3.52}$$

$$\boldsymbol{p}'_n - \boldsymbol{y}_n \to \boldsymbol{0}. \tag{3.53}$$

This implies that

$$\boldsymbol{y}_{n_\ell} \rightharpoonup \widetilde{\boldsymbol{y}}. \tag{3.54}$$

In turn, (3.31) and the relation defining $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ can be rewritten as

$$(\forall n \in \mathbb{N}) \qquad \boldsymbol{u}_n \in \gamma\boldsymbol{A}(\boldsymbol{p}'_n) \tag{3.55}$$

$$\boldsymbol{v}_n \in \gamma\boldsymbol{B}(\boldsymbol{y}_n) \tag{3.56}$$

where

$$(\forall i \in \{1,\dots,m\}) \qquad u_{i,n} = t_{i,n} + \frac{\varepsilon_i}{1-\varepsilon_i}p_{i,n-1} - \frac{1}{1-\varepsilon_i}p'_{i,n} \tag{3.57}$$

$$\boldsymbol{v}_n = 2\boldsymbol{p}_n - \boldsymbol{t}_n - \boldsymbol{y}_n. \tag{3.58}$$

By using (3.10), (3.47), (3.48) and (3.50), it follows that

$$\boldsymbol{u}_{n_\ell} \rightharpoonup \widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{y}} \tag{3.59}$$

while (3.46), (3.49) and (3.54) lead to

$$\boldsymbol{v}_{n_\ell} \rightharpoonup \widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{t}}. \tag{3.60}$$

On the other hand,

$$(\forall n \in \mathbb{N})(\forall i \in \{1,\ldots,m\}) \qquad u_{i,n} + v_{i,n} = p_{i,n} - y_{i,n} + \frac{\varepsilon_i}{1-\varepsilon_i}(p_{i,n-1} - p'_{i,n}) + a_{i,n}. \tag{3.61}$$

According to (3.45), $(\forall i \in \{1,\ldots,k\})$ $p'_{i,n} - p_{i,n-1} \to 0$, and it follows from (3.52) that

$$\boldsymbol{u}_n + \boldsymbol{v}_n \to \boldsymbol{0}. \tag{3.62}$$

In summary, we have built sequences $(\boldsymbol{p}'_n, \boldsymbol{u}_n)_{n\in\mathbb{N}}$ of $\mathrm{gra}(\gamma\boldsymbol{A})$ and $(\boldsymbol{y}_n, \boldsymbol{v}_n)_{n\in\mathbb{N}}$ of $\mathrm{gra}(\gamma\boldsymbol{B})$ satisfying (3.48), (3.53), (3.54), (3.59), (3.60) and (3.62). By invoking now [8, Corollary 3], it can be deduced that

$$\widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{y}} \in \gamma\boldsymbol{A}(\widetilde{\boldsymbol{y}}) \tag{3.63}$$

$$\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{t}} \in \gamma\boldsymbol{B}(\widetilde{\boldsymbol{y}}). \tag{3.64}$$

Summing the two inclusion relations leads to $\widetilde{\boldsymbol{y}} \in \mathrm{zer}(\boldsymbol{A}+\boldsymbol{B})$.
To end the proof, it is sufficient to note that (3.63) is equivalent to $\widetilde{\boldsymbol{y}} = J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}$. This shows that $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ cannot have a weak cluster point other than $J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}$. We have then $\boldsymbol{p}_n \rightharpoonup \widetilde{\boldsymbol{y}}$ and the weak convergence of $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ follows from (3.52). □

Under some restrictive assumptions, a linear convergence property can also be proved:

Proposition 3.6 *Let $\gamma \in]0,+\infty[$, $(\varepsilon_1,\ldots,\varepsilon_m) \in [0,1[^m$ and $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of reals. Let $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{p}_n)_{n\geq -1}$ be sequences generated by Algorithm* (3.30) *when $\boldsymbol{a}_n \equiv \boldsymbol{0}$ and $\boldsymbol{b}_n \equiv \boldsymbol{0}$. Suppose that the following hold.*

(i) $\mathrm{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$

(ii) *There exists $\underline{\lambda} \in]0,2]$ such that $(\forall n \in \mathbb{N})$ $\underline{\lambda} \leq \lambda_{n+1} \leq \lambda_n \leq 2$.*

(iii) *For every $i \in \{1,\ldots,m\}$, $(\varepsilon_i > 0 \Leftrightarrow i \leq k)$.*

(iv) *For every $i \in \{1,\ldots,k\}$, $J_{\gamma(1-\varepsilon_i)A_i}$ is $(1+\tau_{A_i})$-cocoercive and that $R_{\gamma B}$ is $(1+\tau_{\boldsymbol{B}})^{-1/2}$-strictly contractive, where $(\tau_{A_1},\ldots,\tau_{A_k},\tau_{\boldsymbol{B}}) \in]0,+\infty[^{k+1}$ and*

$$(\forall i \in \{1,\ldots,k\}) \qquad \underline{\lambda} > 2\Big(1 - \frac{2\tau_{A_i}}{\varepsilon_i\,\tau_{\boldsymbol{B}}}\Big). \tag{3.65}$$

Then, there exists $\rho \in]0,1[$ such that

$$(\forall n \in \mathbb{N}) \quad |||\boldsymbol{t}_{n+1} - \widetilde{\boldsymbol{t}}|||^2 + \lambda_n \sum_{i=1}^{k} \frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n} - \widetilde{p}_i\|_i^2 \leq \rho^n\Big(|||\boldsymbol{t}_1 - \widetilde{\boldsymbol{t}}|||^2 + \lambda_0 \sum_{i=1}^{k} \frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,0} - \widetilde{p}_i\|_i^2\Big) \tag{3.66}$$

where $(\widetilde{\boldsymbol{t}},\widetilde{\boldsymbol{p}})$ is the unique fixed point of the operator $\boldsymbol{T}$ defined by (3.11)-(3.14).

Proof. Let $(\widetilde{\boldsymbol{t}},\widetilde{\boldsymbol{p}})\in\operatorname{Fix}\boldsymbol{T}$, and set

$$(\forall i\in\{1,\dots,m\})\qquad \widetilde{p}_i'=\begin{cases}\widetilde{p}_i & \text{if } i\le k\\ J_{\gamma A_i}\widetilde{t}_i & \text{otherwise}\end{cases}\tag{3.67}$$

and $(\forall i\in\{k+1,\dots,m\})$ $\tau_{A_i}=0$. Let n be any positive integer. For every $i\in\{1,\dots,m\}$, as a consequence of the cocoercivity of $J_{\gamma(1-\varepsilon_i)A_i}$, we get

$$\begin{aligned}&(1+\tau_{A_i})\|p_{i,n}-\widetilde{p}_i'\|_i^2\le\langle p_{i,n}-\widetilde{p}_i'\mid(1-\varepsilon_i)(t_{i,n}-\widetilde{t}_i)+\varepsilon_i(p_{i,n-1}-\widetilde{p}_i)\rangle_i\\ \Rightarrow\ &-(1-\varepsilon_i)\langle p_{i,n}-\widetilde{p}_i'\mid t_{i,n}-\widetilde{t}\rangle_i\le-\Big(1-\frac{\varepsilon_i}{2}+\tau_{A_i}\Big)\|p_{i,n}-\widetilde{p}_i'\|_i^2+\frac{\varepsilon_i}{2}\|p_{i,n-1}-\widetilde{p}_i\|_i^2.\end{aligned}\tag{3.68}$$

Furthermore, since $R_{\gamma\boldsymbol{B}}$ has been assumed strictly contractive, we have

$$\begin{aligned}&(1+\tau_{\boldsymbol{B}})|||R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n-\boldsymbol{t}_n)-\widetilde{\boldsymbol{t}}|||^2=(1+\tau_{\boldsymbol{B}})|||R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n-\boldsymbol{t}_n)-R_{\gamma\boldsymbol{B}}(2\widetilde{\boldsymbol{p}}'-\widetilde{\boldsymbol{t}})|||^2\\ &\le|||2(\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}')-\boldsymbol{t}_n+\widetilde{\boldsymbol{t}}|||^2=4|||\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}'|||^2+|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2-4\langle\langle\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}'\mid\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}\rangle\rangle.\end{aligned}\tag{3.69}$$

Combined with (3.68), this yields

$$\begin{aligned}&(1+\tau_{\boldsymbol{B}})|||R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n-\boldsymbol{t}_n)-\widetilde{\boldsymbol{t}}|||^2\\ \le\ &-2\sum_{i=1}^m\omega_i\frac{\varepsilon_i+2\tau_{A_i}}{1-\varepsilon_i}\|p_{i,n}-\widetilde{p}_i'\|_i^2+|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+2\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n-1}-\widetilde{p}_i\|_i^2\\ \le\ &-2\sum_{i=1}^k\omega_i\frac{\varepsilon_i+2\tau_{A_i}}{1-\varepsilon_i}\|p_{i,n}-\widetilde{p}_i\|_i^2+|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+2\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n-1}-\widetilde{p}_i\|_i^2.\end{aligned}\tag{3.70}$$

By using (3.37), we have then

$$\begin{aligned}&(1+\tau_{\boldsymbol{B}})|||\boldsymbol{t}_{n+1}-\widetilde{\boldsymbol{t}}|||^2\\ &\le(1+\tau_{\boldsymbol{B}})\Big(\big(1-\frac{\lambda_n}{2}\big)|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\frac{\lambda_n}{2}|||R_{\gamma\boldsymbol{B}}(2\boldsymbol{p}_n-\boldsymbol{t}_n)-\widetilde{\boldsymbol{t}}|||^2\Big)\\ &\le\Big(1+\tau_{\boldsymbol{B}}\big(1-\frac{\lambda_n}{2}\big)\Big)|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\lambda_n\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n-1}-\widetilde{p}_i\|_i^2\\ &\quad-\lambda_n\sum_{i=1}^k\omega_i\frac{\varepsilon_i+2\tau_{A_i}}{1-\varepsilon_i}\|p_{i,n}-\widetilde{p}_i\|_i^2.\end{aligned}\tag{3.71}$$

By using Assumption (ii), it can be deduced that

$$\begin{aligned}&\Big(1+\min\Big\{\tau_{\boldsymbol{B}},\big(\frac{2\tau_{A_i}}{\varepsilon_i}\big)_{1\le i\le k}\Big\}\Big)\Big(|||\boldsymbol{t}_{n+1}-\widetilde{\boldsymbol{t}}|||^2+\lambda_n\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n}-\widetilde{p}_i\|_i^2\Big)\\ &\le\Big(1+\tau_{\boldsymbol{B}}\big(1-\frac{\lambda}{2}\big)\Big)\big(|||\boldsymbol{t}_n-\widetilde{\boldsymbol{t}}|||^2+\lambda_{n-1}\sum_{i=1}^k\frac{\omega_i\varepsilon_i}{1-\varepsilon_i}\|p_{i,n-1}-\widetilde{p}_i\|_i^2\big).\end{aligned}\tag{3.72}$$

By setting now

$$\rho=\frac{1+\tau_{\boldsymbol{B}}\big(1-\frac{\lambda}{2}\big)}{1+\min\{\tau_{\boldsymbol{B}},(2\tau_{A_i}/\varepsilon_i)_{1\le i\le k}\}}\tag{3.73}$$

(3.66) is obtained by induction. The fact that $\rho \in]0,1[$ follows from (3.65). The uniqueness of $(\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{p}})$ is a straightforward consequence of (3.66) and Lemma 3.2(i). $\square$

Remark 3.7 If $(\forall i \in \{1,\dots,k\})$ $A_i = \partial f_i$ where $f_i \in \Gamma_0(\mathcal{H}_i)$ is κ_{f_i}-strongly convex with $\kappa_{f_i} \in]0,+\infty[$, then $J_{\gamma(1-\varepsilon_i)A_i} = \mathrm{prox}_{\gamma(1-\varepsilon_i)f_i}$ is $(1+\tau_{A_i})$-cocoercive with $\tau_{A_i} = \gamma\,(1-\varepsilon_i)\kappa_{f_i}$. (see [15, Proposition 2.5], for example). On the other hand, let $\boldsymbol{B} = \partial \boldsymbol{g}$ where

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) + \frac{\kappa_{\boldsymbol{g}}}{2}|||\boldsymbol{x}|||^2, \tag{3.74}$$

$\boldsymbol{h} \in \Gamma_0(\boldsymbol{\mathcal{H}})$ has a β-Lipschitz continuous gradient and $(\beta, \kappa_{\boldsymbol{g}}) \in]0,+\infty[^2$. Then, it is readily shown that $R_{\gamma\boldsymbol{B}} = \mathrm{rprox}_{\gamma\boldsymbol{g}}$ is $(1+\tau_{\boldsymbol{B}})^{-1/2}$-strictly contractive with $\tau_{\boldsymbol{B}} = \big((1-\gamma\kappa_{\boldsymbol{g}})^2 + 4\gamma\kappa_{\boldsymbol{g}}(1+\gamma\beta(1+\gamma\kappa_{\boldsymbol{g}}))^{-1}\big)^{-1}(1+\gamma\kappa_{\boldsymbol{g}})^2 - 1$ (see the appendix).

4 Zero of a sum of an arbitrary number of maximally monotone operators

Let $(\mathcal{G}, \|\cdot\|)$ be a real Hilbert space and let $\langle \cdot \mid \cdot \rangle$ be the scalar product of $\mathcal{G}$. Let E be a closed linear subspace of $\mathcal{G}$ and let

$$\boldsymbol{L} \colon \mathcal{G} \to \boldsymbol{\mathcal{H}} \colon y \mapsto (L_1 y, \dots, L_m y) \tag{4.1}$$

where, for every $i \in \{1,\dots,m\}$, $L_i \in \mathcal{B}(\mathcal{G}, \mathcal{H}_i)$ is such that $L_i(E)$ is closed. Thus,

$$\big(\forall (\boldsymbol{x}, y) \in \boldsymbol{\mathcal{H}} \times \mathcal{G}\big) \qquad \langle\langle \boldsymbol{x}, \boldsymbol{L}y \rangle\rangle = \sum_{i=1}^m \omega_i \langle x_i, L_i y\rangle_i = \langle \sum_{i=1}^m \omega_i L_i^* x_i, y\rangle. \tag{4.2}$$

This shows that the adjoint of $\boldsymbol{L}$ is

$$\boldsymbol{L}^* \colon \boldsymbol{\mathcal{H}} \to \mathcal{G} \colon (x_1,\dots,x_m) \mapsto \sum_{i=1}^m \omega_i L_i^* x_i. \tag{4.3}$$

In this section, the following problem is considered:

$$\text{find } \ \widetilde{y} \in \mathrm{zer}\Big(\sum_{i=1}^m \omega_i L_i^* \circ A_i \circ L_i + N_E\Big) \tag{4.4}$$

where, for every $i \in \{1,\dots,m\}$, $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ is a maximally monotone operator. Since

$$(\forall y \in \mathcal{G}) \qquad N_E(y) = \begin{cases} E^\perp & y \in E \\ \varnothing & \text{otherwise,} \end{cases} \tag{4.5}$$

Problem (4.4) is equivalent to Problem (1.2).

An algorithm derived from Algorithm (3.30) can be applied to solve Problem (4.4).

Proposition 4.1 *Let* $\gamma \in \,]0, +\infty[$, $(\varepsilon_1, \dots, \varepsilon_m) \in [0, 1[^m$, *and* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence of reals. For every* $i \in \{1, \dots, m\}$, *let* $(a_{i,n})_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}_i$. *Let* $\{y_n\}_{n\in\mathbb{N}} \subset E$, $\{c_n\}_{n\in\mathbb{N}} \subset E$, $\{\boldsymbol{t}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ *and* $\{\boldsymbol{p}_n\}_{n\geq -1} \subset \boldsymbol{\mathcal{H}}$ *be generated by the routine:*

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
(t_{i,0})_{1\leq i\leq m} \in \boldsymbol{\mathcal{H}},\ (p_{i,-1})_{1\leq i\leq m} \in \boldsymbol{\mathcal{H}} \\
y_0 \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,0}\|_i^2
\end{array}
\right. \\
\text{For } n = 0, 1, \dots \\
\left\lfloor
\begin{array}{l}
\text{For } i = 1, \dots, m \\
\quad \lfloor\ p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n} \\
c_n \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2 \\
\text{For } i = 1, \dots, m \\
\quad \lfloor\ t_{i,n+1} = t_{i,n} + \lambda_n\big(L_i(2c_n - y_n) - p_{i,n}\big) \\
y_{n+1} = y_n + \lambda_n(c_n - y_n).
\end{array}
\right.
\end{array}
\tag{4.6}
$$

Suppose that the following assumptions hold.

(i) $\operatorname{zer}\Big(\sum_{i=1}^m \omega_i L_i^* \circ A_i \circ L_i + N_E\Big) \neq \varnothing$

(ii) *There exists* $\underline{\lambda} \in]0, 2[$ *such that* $(\forall n \in \mathbb{N})\ \underline{\lambda} \leq \lambda_{n+1} \leq \lambda_n < 2$.

(iii) $(\forall i \in \{1, \dots, m\})\ \sum_{n\in\mathbb{N}} \|a_{i,n}\|_i < +\infty$.

Then $\big((L_1 y_n, \dots, L_m y_n)\big)_{n\in\mathbb{N}}$, $\big((L_1 c_n, \dots, L_m c_n)\big)_{n\in\mathbb{N}}$, *and* $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ *converge weakly to* $(L_1\widetilde{y}, \dots, L_m\widetilde{y})$ *where* $\widetilde{y}$ *is a solution to Problem* (4.4).

Proof. By using (4.5) and the definition of $\boldsymbol{A}$ in (3.2), we have the following equivalences

$$
\begin{aligned}
& \widetilde{y} \in \operatorname{zer}\Big(\sum_{i=1}^m \omega_i L_i^* \circ A_i \circ L_i + N_E\Big) = \operatorname{zer}(\boldsymbol{L}^* \circ \boldsymbol{A} \circ \boldsymbol{L} + N_E) \\
\Leftrightarrow\quad & \widetilde{y} \in E \quad \text{and} \quad \big(\exists \boldsymbol{z} \in \boldsymbol{A}(\boldsymbol{L}\widetilde{y}), \quad \boldsymbol{L}^*\boldsymbol{z} \in E^\perp\big) \\
\Leftrightarrow\quad & \widetilde{y} \in E \quad \text{and} \quad \big(\exists \boldsymbol{z} \in \boldsymbol{A}(\boldsymbol{L}\widetilde{y}), \quad \boldsymbol{z} \in \boldsymbol{L}(E)^\perp\big) \\
\Leftrightarrow\quad & \widetilde{y} \in E \quad \text{and} \quad \boldsymbol{L}\widetilde{y} \in \operatorname{zer}(\boldsymbol{A} + N_{\boldsymbol{D}})
\end{aligned}
\tag{4.7}
$$

where $\boldsymbol{D} = \boldsymbol{L}(E)$. Hence, Problem (4.4) reduces to finding a zero of $\boldsymbol{A} + N_{\boldsymbol{D}}$. This latter problem is a specialization of Problem (3.1) when $\boldsymbol{B} = N_{\boldsymbol{D}}$. We have then $J_{\gamma\boldsymbol{B}} = J_{\gamma\partial\iota_{\boldsymbol{D}}} = \operatorname{prox}_{\iota_{\boldsymbol{D}}} = P_{\boldsymbol{D}}$. According to Proposition 3.3, under Assumptions (i)-(iii), the algorithm:

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
\boldsymbol{t}_0 \in \boldsymbol{\mathcal{H}}, \boldsymbol{p}_{-1} \in \boldsymbol{\mathcal{H}}
\end{array}
\right. \\
\text{For } n = 0, 1, \dots \\
\left\lfloor
\begin{array}{l}
\text{For } i = 1, \dots, m \\
\quad \lfloor\ p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n} \\
\boldsymbol{d}_n = P_{\boldsymbol{D}}\, \boldsymbol{p}_n \\
\boldsymbol{x}_n = P_{\boldsymbol{D}}\, \boldsymbol{t}_n \\
\boldsymbol{t}_{n+1} = \boldsymbol{t}_n + \lambda_n\Big(2\boldsymbol{d}_n - \boldsymbol{x}_n - \boldsymbol{p}_n\Big)
\end{array}
\right.
\end{array}
\tag{4.8}
$$

allows us to generate a sequence $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ converging weakly to $\widetilde{\boldsymbol{t}}$ such that $J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}} \in \operatorname{zer}(\boldsymbol{A} + N_{\boldsymbol{D}})$. As a consequence,

$$J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}} \in \boldsymbol{D}. \tag{4.9}$$

According to the fixed point properties of the Douglas-Rachford algorithm,

$$\widetilde{\boldsymbol{t}} = (2P_{\boldsymbol{D}} - \mathrm{Id}\,)(2J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{t}}). \tag{4.10}$$

Due to (4.9), (4.10) is equivalent to

$$P_{\boldsymbol{D}}\widetilde{\boldsymbol{t}} = J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}. \tag{4.11}$$

By using the weak continuity of $P_{\boldsymbol{D}}$, it can be deduced that

$$\boldsymbol{x}_n = P_{\boldsymbol{D}}\boldsymbol{t}_n \rightharpoonup P_{\boldsymbol{D}}\widetilde{\boldsymbol{t}} = J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}. \tag{4.12}$$

In other words, $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly to $\boldsymbol{L}\widetilde{y} = J_{\gamma\boldsymbol{A}}\widetilde{\boldsymbol{t}}$ where $\widetilde{y} \in E$ is a solution to Problem (4.4).

In addition, $(\boldsymbol{x}_n)_{n>0}$ can be computed in a recursive manner through the relation

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_{n+1} = P_{\boldsymbol{D}}\boldsymbol{t}_{n+1} = P_{\boldsymbol{D}}\boldsymbol{t}_n + \lambda_n(2\boldsymbol{d}_n - \boldsymbol{x}_n - P_{\boldsymbol{D}}\boldsymbol{p}_n) = \boldsymbol{x}_n + \lambda_n(\boldsymbol{d}_n - \boldsymbol{x}_n) \tag{4.13}$$

where we have used the fact that $(\boldsymbol{x}_n)_{n\geq 0}$ and $(\boldsymbol{d}_n)_{n\geq 0}$ are sequences of $\boldsymbol{D}$. Algorithm (4.8) then becomes

$$\begin{array}{l}
\text{Initialization} \\
\left\lfloor \begin{array}{l} \boldsymbol{t}_0 \in \boldsymbol{\mathcal{H}}, \boldsymbol{p}_{-1} \in \boldsymbol{\mathcal{H}} \\ \boldsymbol{x}_0 = P_{\boldsymbol{D}}\boldsymbol{t}_0 \end{array} \right. \\
\text{For } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l} \text{For } i = 1, \ldots, m \\ \quad \lfloor\; p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n} \\ \boldsymbol{d}_n = P_{\boldsymbol{D}}\,\boldsymbol{p}_n \\ \boldsymbol{t}_{n+1} = \boldsymbol{t}_n + \lambda_n\Big(2\boldsymbol{d}_n - \boldsymbol{x}_n - \boldsymbol{p}_n\Big) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{d}_n - \boldsymbol{x}_n). \end{array} \right.
\end{array} \tag{4.14}$$

According to Assumption (ii), $(\lambda_n)_{n\in\mathbb{N}}$ converges to a positive limit, and we deduce from (4.13) that $\boldsymbol{d}_n \rightharpoonup \boldsymbol{L}\widetilde{y}$. In turn, the last update equation in Algorithm (4.8) yields $\boldsymbol{p}_n \rightharpoonup \boldsymbol{L}\widetilde{y}$.

Finally, it can be noticed that, for every $\boldsymbol{u} \in \boldsymbol{\mathcal{H}}$, $\boldsymbol{v} = P_{\boldsymbol{D}}\boldsymbol{u}$ if and only if $\boldsymbol{v} = \boldsymbol{L}w$ where

$$w \in \operatorname{Argmin}_{z\in E} |||\boldsymbol{L}z - \boldsymbol{u}|||. \tag{4.15}$$

Since by construction $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$ are sequences in $\boldsymbol{D}$, for every $n \in \mathbb{N}$, there exists $y_n \in E$ and $c_n \in E$ such that $\boldsymbol{x}_n = \boldsymbol{L}y_n$ and $\boldsymbol{d}_n = \boldsymbol{L}c_n$. Hence, Algorithm (4.6) appears as an implementation of Algorithm (4.14). □

Algorithm (4.14) requires to compute the linear projection onto $\boldsymbol{D} = \boldsymbol{L}(E)$ which amounts to solving the quadratic programming problem (4.15) for any $\boldsymbol{u} \in \boldsymbol{\mathcal{H}}$. Note that $w \in \mathcal{G}$ is a solution to (4.15) if and only if

$$\big(w, \boldsymbol{L}^*(\boldsymbol{L}w - \boldsymbol{u})\big) \in E \times E^{\perp}. \tag{4.16}$$

This shows that this problem can also be formulated in terms of the generalized inverse of $\boldsymbol{L}$ w.r.t. E. In particular, if $\boldsymbol{L}^*\boldsymbol{L} = \sum_{i=1}^m \omega_i L_i^* L_i$ is an isomorphism on $\mathcal{G}$, for every $\boldsymbol{u} \in \mathcal{H}$, there exists a unique $w \in \mathcal{G}$ satisfying (4.16) (since $z \mapsto |||\boldsymbol{L}z - \boldsymbol{u}|||^2$ is strictly convex).

In this case, a stronger convergence result can be obtained as stated below.

Corollary 4.2 *Suppose that the assumptions of Proposition 4.1 hold and that* $\sum_{i=1}^m \omega_i L_i^* L_i$ *is an isomorphism. Let* $(y_n)_{n\in\mathbb{N}}$ *be a sequence generated by Algorithm* (4.6)*. Then* $(y_n)_{n\in\mathbb{N}}$ *converges weakly to a solution to Problem* (4.4)*.*

Proof. According to Proposition 4.1,

$$\boldsymbol{L} y_n \rightharpoonup \boldsymbol{L}\widetilde{y} \tag{4.17}$$

where $\widetilde{y}$ is a solution to Problem (4.4). By recalling that any bounded linear operator is weakly continuous, we have therefore

$$\boldsymbol{L}^*\boldsymbol{L} y_n \rightharpoonup \boldsymbol{L}^*\boldsymbol{L}\widetilde{y} \tag{4.18}$$

and, consequently,

$$y_n = (\boldsymbol{L}^*\boldsymbol{L})^{-1}\boldsymbol{L}^*\boldsymbol{L} y_n \rightharpoonup (\boldsymbol{L}^*\boldsymbol{L})^{-1}\boldsymbol{L}^*\boldsymbol{L}\widetilde{y} = \widetilde{y}. \tag{4.19}$$

□

The assumption that $\sum_{i=1}^m \omega_i L_i^* L_i$ is an isomorphism may be restrictive. However, a variant of Algorithm (4.6) allows us to relax this requirement.

Proposition 4.3 *Let* $\gamma \in \,]0,+\infty[$, $(\varepsilon_1,\ldots,\varepsilon_m) \in [0,1[^m$, $\alpha \in \,]0,+\infty[$, *and* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence of reals. For every* $i \in \{1,\ldots,m\}$, *let* $(a_{i,n})_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}_i$*. Let* $\{y_n\}_{n\in\mathbb{N}} \subset E$, $\{c_n\}_{n\in\mathbb{N}} \subset E$, $\{r_n\}_{n\in\mathbb{N}} \subset \mathcal{G}$, $\{\boldsymbol{t}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$, *and* $\{\boldsymbol{p}_n\}_{n\geq -1} \subset \boldsymbol{\mathcal{H}}$ *be generated by the routine:*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
(t_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\ (p_{i,-1})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\ r_0 \in \mathcal{G}\\
y_0 = \arg\min_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,0}\|_i^2 + \alpha\|z - r_0\|^2
\end{array}\right.\\
\text{For } n = 0,1,\ldots\\
\left\lfloor\begin{array}{l}
\text{For } i = 1,\ldots,m\\
\quad\lfloor\ p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n}\\
c_n = \arg\min_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2 + \alpha\|z - r_n\|^2\\
\text{For } i = 1,\ldots,m\\
\quad\lfloor\ t_{i,n+1} = t_{i,n} + \lambda_n\big(L_i(2c_n - y_n) - p_{i,n}\big)\\
r_{n+1} = r_n + \lambda_n(2c_n - y_n - r_n)\\
y_{n+1} = y_n + \lambda_n(c_n - y_n).
\end{array}\right.
\end{array} \tag{4.20}$$

Suppose that Assumptions (i)-(iii) *in Proposition 4.1 hold. Then,* $(y_n)_{n\in\mathbb{N}}$ *converges weakly to a solution to Problem* (4.4)*.*

Proof. Problem (4.4) can be reformulated as

$$\text{find } \widetilde{y} \in \operatorname{zer}\Big(\sum_{i=1}^{m+1} \omega_i L_i^* \circ A_i \circ L_i + N_E\Big) \quad \text{where } A_{m+1} = 0 \text{ and } L_{m+1} = \mathrm{Id}\,. \tag{4.21}$$

In this case, for every $\omega_{m+1} \in \,]0,+\infty[$, the self-adjoint operator $\sum_{i=1}^{m+1} \omega_i L_i^* L_i = \sum_{i=1}^m \omega_i L_i^* L_i + \omega_{m+1}\,\mathrm{Id}$ is an isomorphism. By applying a version of Algorithm (4.6) to the above problem with

$m+1$ operators, we get:

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor \begin{array}{l}
(t_{i,0})_{1\le i\le m+1} \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_m \times \mathcal{G},\ (p_{i,-1})_{1\le i\le m+1} \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_m \times \mathcal{G} \\
y_0 \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,0}\|_i^2 + \omega_{m+1}\|z - t_{m+1,0}\|^2
\end{array}\right. \\
\text{For } n = 0,1,\ldots \\
\left\lfloor \begin{array}{l}
\text{For } i = 1,\ldots,m+1 \\
\quad \left\lfloor\ p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n} \right. \\
c_n \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2 + \omega_{m+1}\|z - p_{m+1,n}\|^2 \\
\text{For } i = 1,\ldots,m+1 \\
\quad \left\lfloor\ t_{i,n+1} = t_{i,n} + \lambda_n\big(L_i(2c_n - y_n) - p_{i,n}\big) \right. \\
y_{n+1} = y_n + \lambda_n(c_n - y_n).
\end{array}\right.
\end{array}
\tag{4.22}
$$

Setting $\alpha = \omega_{m+1}$, defining $(\forall n \in \mathbb{N})\ r_n = t_{m+1,n}$, and setting also $\varepsilon_{m+1} = 0$ and $a_{m+1,n} \equiv 0$ yields $p_{m+1,n} \equiv t_{m+1,n}$, and Algorithm (4.20) is derived. $\square$

Remark 4.4 Another variant of Algorithm (4.6) which allows us to take into account errors in the computation of the projection onto $\boldsymbol{D}$ is the following:

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor\ (t_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\ (p_{i,-1})_{1\le i\le m} \in \boldsymbol{\mathcal{H}} \right. \\
\text{For } n = 0,1,\ldots \\
\left\lfloor \begin{array}{l}
\text{For } i = 1,\ldots,m \\
\quad \left\lfloor\ p_{i,n} = J_{\gamma(1-\varepsilon_i)A_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n} \right. \\
c_n - b_n^c \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2 \\
y_n - b_n^y \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,n}\|_i^2 \\
\boldsymbol{t}_{n+1} = \boldsymbol{t}_n + \lambda_n\big(\boldsymbol{L}(2c_n - y_n) - \boldsymbol{p}_n\big).
\end{array}\right.
\end{array}
\tag{4.23}
$$

Under the assumptions of Proposition 4.1 and provided that the error sequences $(b_n^y)_{n\in\mathbb{N}}$ and $(b_n^c)_{n\in\mathbb{N}}$ are such that

$$
\sum_{n=0}^{+\infty} \|b_n^y\| + \|b_n^c\| < +\infty \tag{4.24}
$$

the same convergence results as for Proposition 4.1 can be proved. Note however that this algorithm requires two projections onto $\boldsymbol{L}(E)$ at each iteration. Algorithm (34) in [18] is a special case of this algorithm corresponding to the case when $\varepsilon_1 = \ldots = \varepsilon_m = 0$, $\mathcal{H}_1 = \ldots = \mathcal{H}_m = E = \mathcal{G}$, $L_1 = \ldots = L_m = \operatorname{Id}$, and $b_n^y \equiv b_n^c \equiv 0$.

5 Application to convex optimization

In this section, the following optimization problem is considered:

$$
\underset{y\in E}{\text{minimize}} \ \sum_{i=1}^m f_i(L_i y) \tag{5.1}
$$

where, for every $i \in \{1,\dots,m\}$, $f_i \in \Gamma_0(\mathcal{H}_i)$. As a preliminary result, it can be noticed that:

Proposition 5.1 *If there exists $\overline{y} \in E$ such that*

$$(\forall i \in \{1,\dots,m\}) \qquad L_i\overline{y} \in \operatorname{dom} f_i \tag{5.2}$$

and $\lim_{y\in E, \|L_1 y\|_1+\cdots+\|L_m y\|_m \to +\infty} f_1(L_1 y) + \cdots + f_m(L_m y) = +\infty$, *then the set of solutions to Problem* (5.1) *is nonempty.*

Proof. Problem (5.1) is equivalent to

$$\underset{\boldsymbol{x}\in\boldsymbol{D}}{\text{minimize}} \;\; \boldsymbol{f}(\boldsymbol{x}) \quad \text{where} \quad \boldsymbol{D} = \boldsymbol{L}(E). \tag{5.3}$$

As a consequence of classical results of convex analysis [28, Proposition II.1.2], the set of solutions to Problem (5.3) is nonempty if $\operatorname{dom}\boldsymbol{f} \cap \boldsymbol{D} = \operatorname{dom}\boldsymbol{f} \cap \boldsymbol{L}(E) \neq \varnothing$ and

$$\lim_{\boldsymbol{x}\in\boldsymbol{D},\, |||\boldsymbol{x}|||\to+\infty} \boldsymbol{f}(\boldsymbol{x}) = +\infty \tag{5.4}$$

which yields the desired result. □

An algorithm derived from Algorithm (4.6) can be applied to solve Problem (5.1). In the following, $(\omega_i)_{1\le i\le m}$ are positive constants, as in the previous section.

Proposition 5.2 *Let $(\varepsilon_1,\dots,\varepsilon_m) \in [0,1[^m$ and $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of reals. For every $i \in \{1,\dots,m\}$, let $(a_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}_i$. Let $\{y_n\}_{n\in\mathbb{N}} \subset E$, $\{c_n\}_{n\in\mathbb{N}} \subset E$, $\{\boldsymbol{t}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ and $\{\boldsymbol{p}_n\}_{n\ge-1} \subset \boldsymbol{\mathcal{H}}$ be generated by the routine:*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
(t_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\ (p_{i,-1})_{1\le i\le m} \in \boldsymbol{\mathcal{H}}\\
y_0 \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,0}\|_i^2
\end{array}\right.\\
\text{For } n = 0,1,\dots\\
\left\lfloor\begin{array}{l}
\text{For } i = 1,\dots,m\\
\left\lfloor\begin{array}{l}
p_{i,n} = \operatorname{prox}_{\frac{(1-\varepsilon_i)f_i}{\omega_i}}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n}
\end{array}\right.\\
c_n \in \operatorname{Arg\,min}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2\\
\text{For } i = 1,\dots,m\\
\left\lfloor\begin{array}{l}
t_{i,n+1} = t_{i,n} + \lambda_n\big(L_i(2c_n - y_n) - p_{i,n}\big)
\end{array}\right.\\
y_{n+1} = y_n + \lambda_n(c_n - y_n).
\end{array}\right.
\end{array} \tag{5.5}$$

Suppose that the following assumptions hold.

(i) $\boldsymbol{0} \in \operatorname{sri}\big\{(L_1 z - x_1,\dots,L_m z - x_m) \;\big|\; z\in E,\, x_1 \in \operatorname{dom} f_1,\dots,\, x_m \in \operatorname{dom} f_m\big\}$.

(ii) *There exists $\underline{\lambda} \in]0,2[$ such that $(\forall n\in\mathbb{N})$ $\underline{\lambda} \le \lambda_{n+1} \le \lambda_n < 2$.*

(iii) $(\forall i \in \{1,\dots,m\})$ $\sum_{n\in\mathbb{N}} \|a_{i,n}\|_i < +\infty$.

If the set of solutions to Problem (5.1) *is nonempty, then $\big((L_1 y_n,\dots,L_m y_n)\big)_{n\in\mathbb{N}}$, $\big((L_1 c_n,\dots,L_m c_n)\big)_{n\in\mathbb{N}}$, and $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ converge weakly to $(L_1\widetilde{y},\dots,L_m\widetilde{y})$ where $\widetilde{y}$ is a solution to Problem* (5.1).

Proof. Problem (5.1) is equivalent to

$$\underset{y\in\mathcal{G}}{\text{minimize}} \;\; \boldsymbol{f}(\boldsymbol{L}y) + \iota_E(y). \tag{5.6}$$

A necessary and sufficient condition [53, Theorem 2.5.7] for $\widetilde{y}$ to be a solution to the above problem is:

$$\widetilde{y} \in \operatorname{zer} \partial(\boldsymbol{f} \circ \boldsymbol{L} + \iota_E). \tag{5.7}$$

Provided that

$$\boldsymbol{0} \in \operatorname{sri} \big\{ \boldsymbol{L}z - \boldsymbol{x} \;\big|\; z \in E,\, \boldsymbol{x} \in \operatorname{dom} \boldsymbol{f} \big\} \tag{5.8}$$

we have [53, Theorem 2.8.3]

$$\partial(\boldsymbol{f} \circ \boldsymbol{L} + \iota_E) = \boldsymbol{L}^* \circ \partial \boldsymbol{f} \circ \boldsymbol{L} + \partial\iota_E. \tag{5.9}$$

Due Assumption (i), (5.8) is obviously satisfied and (5.7) is therefore equivalent to

$$\widetilde{y} \in \operatorname{zer}(\boldsymbol{L}^* \circ \partial \boldsymbol{f} \circ \boldsymbol{L} + \partial\iota_E) = \operatorname{zer}\Big(\sum_{i=1}^m L_i^* \circ \partial f_i \circ L_i + N_E\Big) \tag{5.10}$$

where we have used the fact that $\partial \boldsymbol{f} = (\omega_1^{-1}\partial f_1) \times \cdots \times (\omega_m^{-1}\partial f_m)$, which follows from the definition of the norm in (2.1). So, Problem 5.1 appears as a specialisation of Problem 4.4 when $(\forall i \in \{1,\ldots,m\})\; A_i = \partial f_i/\omega_i$. Algorithm (5.5) is then derived from Algorithm (4.6) (with $\gamma = 1$) since $J_{(1-\varepsilon_i)A_i} = \operatorname{prox}_{(1-\varepsilon_i)f_i/\omega_i}$ and its convergence follows from Proposition 4.1. ☐

When $\sum_{i=1}^m \omega_i L_i^* L_i$ is an isomorphism on $\mathcal{G}$, a stronger convergence result can be deduced from Corollary 4.2.

Corollary 5.3 *Suppose that the assumptions of Proposition 5.2 hold and that $\sum_{i=1}^m \omega_i L_i^* L_i$ is an isomorphism. Let $(y_n)_{n\in\mathbb{N}}$ be a sequence generated by Algorithm* (5.5). *If the set of solutions to Problem* (5.1) *is nonempty, then $(y_n)_{n\in\mathbb{N}}$ converges weakly to an element of this set.*

A variant of Algorithm (5.5) is applicable when $\sum_{i=1}^m \omega_i L_i^* L_i$ is not an isomorphism.

Proposition 5.4 *Let $(\varepsilon_1,\ldots,\varepsilon_m) \in [0,1[^m$, $\alpha \in]0,+\infty[$, and $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of reals. For every $i \in \{1,\ldots,m\}$, let $(a_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}_i$. Let $\{y_n\}_{n\in\mathbb{N}} \subset E$, $\{c_n\}_{n\in\mathbb{N}} \subset E$, $\{r_n\}_{n\in\mathbb{N}} \subset \mathcal{G}$ $\{\boldsymbol{t}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ and $\{\boldsymbol{p}_n\}_{n\geq -1} \subset \boldsymbol{\mathcal{H}}$ be generated by the routine:*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
(t_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\; (p_{i,-1})_{1\le i\le m} \in \boldsymbol{\mathcal{H}},\; r_0 \in \mathcal{G}\\
y_0 = \arg\min_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - t_{i,0}\|_i^2 + \alpha\|z - r_0\|^2
\end{array}\right.\\
\text{For } n = 0,1,\ldots\\
\left\lfloor\begin{array}{l}
\text{For } i = 1,\ldots,m\\
\left\lfloor\begin{array}{l}
p_{i,n} = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i}\big((1-\varepsilon_i)t_{i,n} + \varepsilon_i p_{i,n-1}\big) + a_{i,n}
\end{array}\right.\\
c_n = \arg\min_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - p_{i,n}\|_i^2 + \alpha\|z - r_n\|^2\\
\text{For } i = 1,\ldots,m\\
\left\lfloor\begin{array}{l}
t_{i,n+1} = t_{i,n} + \lambda_n\big(L_i(2c_n - y_n) - p_{i,n}\big)
\end{array}\right.\\
r_{n+1} = r_n + \lambda_n(2c_n - y_n - r_n)\\
y_{n+1} = y_n + \lambda_n(c_n - y_n).
\end{array}\right.
\end{array} \tag{5.11}$$

Suppose that Assumptions (i)-(iii) *in Proposition 5.2 hold. If the set of solutions to Problem* (5.1) *is nonempty, then* $(y_n)_{n\in\mathbb{N}}$ *converges weakly to an element of this set.*

Proof.
By proceeding similarly to the proof of Proposition 5.2, it can be deduced from Proposition 4.3 that the sequence $(y_n)_{n\in\mathbb{N}}$ generated by Algorithm (5.11) converges to a solution to Problem (5.1) when such a solution exists. □

Remark 5.5

(i) When $\boldsymbol{\mathcal{H}}$ is finite dimensional, Assumptions (i) in Proposition 5.2 takes the form

$$\exists \overline{y} \in E, \qquad L_1\overline{y} \in \operatorname{ri}\operatorname{dom} f_1, \dots, L_m\overline{y} \in \operatorname{ri}\operatorname{dom} f_m. \tag{5.12}$$

Indeed, in this case, we have (see Proposition 3.6 in [21]):

$$\begin{aligned} & \mathbf{0} \in \operatorname{sri}\{\boldsymbol{D} - \operatorname{dom}\boldsymbol{f}\} = \operatorname{ri}\{\boldsymbol{D} - \operatorname{dom}\boldsymbol{f}\} = \operatorname{ri}\boldsymbol{D} - \operatorname{ri}\operatorname{dom}\boldsymbol{f} = \boldsymbol{D} - \operatorname{ri}\operatorname{dom}\boldsymbol{f} \\ \Leftrightarrow\quad & \boldsymbol{D} \cap \operatorname{ri}\operatorname{dom}\boldsymbol{f} \neq \varnothing. \end{aligned} \tag{5.13}$$

(ii) We have seen that Problem (5.1) can be put under the form of Problem (5.3). The dual formulation of this problem is

$$\underset{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}}{\operatorname{minimize}}\ \iota_{\boldsymbol{D}}^*(-\boldsymbol{x}) + \boldsymbol{f}^*(\boldsymbol{x}). \tag{5.14}$$

Recall that $\iota_{\boldsymbol{D}}^* = \iota_{\boldsymbol{D}^\perp}$. The dual problem can thus be rewritten as

$$\underset{\boldsymbol{x}\in\boldsymbol{D}^\perp}{\operatorname{minimize}}\ \boldsymbol{f}^*(\boldsymbol{x}). \tag{5.15}$$

Algorithm (4.8) can again be used to solve the dual problem. However, since $P_{\boldsymbol{D}^\perp} = \operatorname{Id} - P_{\boldsymbol{D}}$ and, $(\forall \gamma > 0)$ $(\forall i \in \{1,\dots,m\})$ $\operatorname{prox}_{\gamma f_i^*} = \operatorname{Id} - \gamma \operatorname{prox}_{f_i/\gamma}(\cdot/\gamma)$ [23, Lemma 2.10], the resulting algorithm for the dual problem takes a form very similar to the algorithm proposed for the primal one.

6 Connections with some other parallel splitting optimization algorithms

Firstly, is interesting to note that PPXA (Algorithm 3.1 in [21]) is a special case of Algorithm (5.5) corresponding to the case when $\varepsilon_1 = \ldots = \varepsilon_m = 0$, $\mathcal{H}_1 = \ldots = \mathcal{H}_m = E = \mathcal{G}$, and $L_1 = \ldots = L_m = \operatorname{Id}$.

In order to better emphasize the link existing with other algorithms, it will appear useful to rewrite Algorithm (5.5) in a different form.

Proposition 6.1 *Let* $(\varepsilon_1,\dots,\varepsilon_m)\in[0,1[^m$ *and* $(\lambda_n)_{n\geq -2}$ *be a sequence of reals. Algorithm* (5.5) *with* $(p_{i,-1})_{1\leq i\leq m}=(L_1y_0,\dots,L_my_0)$ *is equivalent to the following routine:*

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\lambda_{-2}=\lambda_{-1}=1\\
(e_{i,0})_{1\leq i\leq m}\in\boldsymbol{\mathcal{H}}\\
u_0\in\operatorname{Argmin}_{z\in E}\sum_{i=1}^m\omega_i\|L_iz+e_{i,0}\|_i^2\\
\text{For } i=1,\dots,m\\
\left\lfloor
\begin{array}{l}
\ell_{i,0}=0\\
w_{i,-1}=0\\
w_{i,0}=L_iu_{i,0}
\end{array}
\right.
\end{array}
\right.\\
\text{For } n=0,1,\dots\\
\left\lfloor
\begin{array}{l}
u_{n+1}\in\operatorname{Argmin}_{z\in E}\sum_{i=1}^m\omega_i\|L_iz-\lambda_{n-1}w_{i,n}+e_{i,n}\|_i^2\\
\text{For } i=1,\dots,m\\
\left\lfloor
\begin{array}{l}
v_{i,n+1}=L_iu_{n+1}\\
w_{i,n+1}=\operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i}f_i}\big((1-\varepsilon_i)(v_{i,n+1}+e_{i,n}-\ell_{i,n})+\varepsilon_iw_{i,n}\big)+a_{i,n}
\end{array}
\right.\\
k_{n+1}\in\operatorname{Argmin}_{z\in E}\sum_{i=1}^m\omega_i\|L_iz-e_{i,n}+\lambda_{n-2}w_{i,n-1}\|_i^2\\
\text{For } i=1,\dots,m\\
\left\lfloor
\begin{array}{l}
\ell_{i,n+1}=(1-\lambda_n)L_ik_{n+1}\\
e_{i,n+1}=v_{i,n+1}+e_{i,n}-\lambda_nw_{i,n+1}+\ell_{i,n+1}.
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{6.1}
$$

Suppose that Assumptions (i), (ii) *and* (iii) *in Proposition 5.2 hold. If the set of solutions to Problem* (5.1) *is nonempty, then* $\big((v_{1,n},\dots,v_{m,n})\big)_{n>0}$ *converges weakly to* $(L_1\widetilde{y},\dots,L_m\widetilde{y})$ *where* $\widetilde{y}$ *is a solution to Problem* (5.1).

Proof. Let us start from Algorithm (4.8) and express it as follows: for every $n\in\mathbb{N}$,

$$
\left\{
\begin{array}{l}
(\forall i\in\{1,\dots,m\})\ p_{i,n}^D+p_{i,n}^\perp=\operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i}f_i}\big((1-\varepsilon_i)(t_{i,n}^D+t_{i,n}^\perp)+\varepsilon_i(p_{i,n-1}^D+p_{i,n-1}^\perp)\big)+a_{i,n}\\
\boldsymbol{d}_n=\boldsymbol{p}_n^D\\
\boldsymbol{x}_n=\boldsymbol{t}_n^D\\
\boldsymbol{t}_{n+1}^D=\boldsymbol{t}_n^D+\lambda_n\big(2\boldsymbol{d}_n-\boldsymbol{x}_n-\boldsymbol{p}_n^D\big)\\
\boldsymbol{t}_{n+1}^\perp=\boldsymbol{t}_n^\perp-\lambda_n\boldsymbol{p}_n^\perp
\end{array}
\right.
\tag{6.2}
$$

where $\boldsymbol{p}_n^D$ (resp $\boldsymbol{p}_n^\perp$) denotes the projection of $\boldsymbol{p}_n$ onto $\boldsymbol{D}$ (resp. $\boldsymbol{D}^\perp$), and $p_{i,n}^D$ (resp. $p_{i,n}^\perp$) is its i-th component in $\mathcal{H}_i$, a similar notation being used for the other variables. The above set of equations can be rewritten as

$$
\left\{
\begin{array}{l}
(\forall i\in\{1,\dots,m\})\ p_{i,n}^D+p_{i,n}^\perp=\operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i}f_i}\big((1-\varepsilon_i)(t_{i,n}^D+t_{i,n}^\perp)+\varepsilon_i(p_{i,n-1}^D+p_{i,n-1}^\perp)\big)+a_{i,n}\\
\boldsymbol{t}_{n+1}^D=\boldsymbol{t}_n^D+\lambda_n\big(\boldsymbol{p}_n^D-\boldsymbol{t}_n^D\big)\\
\boldsymbol{t}_{n+1}^\perp=\boldsymbol{t}_n^\perp-\lambda_n\boldsymbol{p}_n^\perp.
\end{array}
\right.
\tag{6.3}
$$

Let us now introduce sequences $(\boldsymbol{w}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{e}_n)_{n\in\mathbb{N}}$ such that

$$\boldsymbol{w}_n=\boldsymbol{p}_{n-1} \tag{6.4}$$

$$\boldsymbol{e}_n^\perp=\boldsymbol{t}_n^\perp. \tag{6.5}$$

Then, for every $n \in \mathbb{N}$, (6.3) becomes

$$\begin{cases} (\forall i \in \{1,\dots,m\}) \\ w_{i,n+1}^D + w_{i,n+1}^\perp = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i} \Big((1-\varepsilon_i)\big((1-\lambda_{n-1}) t_{i,n-1}^D + \lambda_{n-1} w_{i,n}^D + e_{i,n}^\perp\big) + \varepsilon_i (w_{i,n}^D + w_{i,n}^\perp) \Big) \\ \qquad\qquad +a_{i,n} \\ \boldsymbol{t}_{n+1}^D = (1-\lambda_n)\boldsymbol{t}_n^D + \lambda_n \boldsymbol{w}_{n+1}^D \\ \boldsymbol{e}_{n+1}^\perp = \boldsymbol{e}_n^\perp - \lambda_n \boldsymbol{w}_{n+1}^\perp \end{cases} \tag{6.6}$$

provided that $\lambda_{-1} = 1$ (thus allowing us to choose $t_{i,-1}^D$ arbitrarily) since $\boldsymbol{p}_{-1} = (L_1 y_0, \dots, L_m y_0) = \boldsymbol{x}_0 \in \boldsymbol{D} \Leftrightarrow \boldsymbol{w}_0^D = \boldsymbol{p}_{-1} = \boldsymbol{x}_0 = \boldsymbol{t}_0^D$. In addition, for every $n \in \mathbb{N}$, set

$$\begin{aligned} \boldsymbol{e}_{n+1}^D &= \lambda_{n-1} \boldsymbol{p}_{n-1}^D - \boldsymbol{t}_{n+1}^D \\ &= \lambda_{n-1} \boldsymbol{w}_n^D - \boldsymbol{t}_{n+1}^D. \end{aligned} \tag{6.7}$$

Then, we have: for every $n \in \mathbb{N}$,

$$\begin{aligned} & \boldsymbol{t}_{n+1}^D = (1-\lambda_n)\boldsymbol{t}_n^D + \lambda_n \boldsymbol{w}_{n+1}^D \\ \Leftrightarrow \quad & \boldsymbol{e}_{n+1}^D = (1-\lambda_n)(\boldsymbol{e}_n^D - \lambda_{n-2}\boldsymbol{w}_{n-1}^D) + \lambda_{n-1}\boldsymbol{w}_n^D - \lambda_n \boldsymbol{w}_{n+1}^D \end{aligned} \tag{6.8}$$

provided that $\lambda_{-2} = 1$ and $\boldsymbol{w}_{-1}^D - \boldsymbol{e}_0^D = \boldsymbol{w}_0^D$. By using the two previous relations, we see that (6.6) is equivalent to

$$\begin{cases} (\forall i \in \{1,\dots,m\}) \\ w_{i,n+1}^D + w_{i,n+1}^\perp = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i} \Big((1-\varepsilon_i)\big((1-\lambda_{n-1})(\lambda_{n-3} w_{i,n-2}^D - e_{i,n-1}^D) + \lambda_{n-1} w_{i,n}^D + e_{i,n}^\perp\big) \\ \qquad\qquad +\varepsilon_i (w_{i,n}^D + w_{i,n}^\perp) \Big) + a_{i,n} \\ \boldsymbol{e}_{n+1}^D = (1-\lambda_n)(\boldsymbol{e}_n^D - \lambda_{n-2}\boldsymbol{w}_{n-1}^D) + \lambda_{n-1}\boldsymbol{w}_n^D - \lambda_n \boldsymbol{w}_{n+1}^D \\ \boldsymbol{e}_{n+1}^\perp = \boldsymbol{e}_n^\perp - \lambda_n \boldsymbol{w}_{n+1}^\perp. \end{cases} \tag{6.9}$$

By introducing intermediate variables $\boldsymbol{v}_{n+1}$ and $\boldsymbol{\ell}_{n+1}$, this can be rewritten as

$$\begin{cases} \boldsymbol{v}_{n+1} = \lambda_{n-1}\boldsymbol{w}_n^D - \boldsymbol{e}_n^D \\ (\forall i \in \{1,\dots,m\}) \quad w_{i,n+1}^D + w_{i,n+1}^\perp \\ \qquad = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i} \big((1-\varepsilon_i)(v_{i,n+1} + e_{i,n}^D + e_{i,n}^\perp - \ell_{i,n}) + \varepsilon_i (w_{i,n}^D + w_{i,n}^\perp) \big) + a_{i,n} \\ \boldsymbol{\ell}_{n+1} = (1-\lambda_n)(\boldsymbol{e}_n^D - \lambda_{n-2}\boldsymbol{w}_{n-1}^D) \\ \boldsymbol{e}_{n+1}^D = \boldsymbol{v}_{n+1} + \boldsymbol{e}_n^D - \lambda_n \boldsymbol{w}_{n+1}^D + \boldsymbol{\ell}_{n+1} \\ \boldsymbol{e}_{n+1}^\perp = \boldsymbol{e}_n^\perp - \lambda_n \boldsymbol{w}_{n+1}^\perp \end{cases} \tag{6.10}$$

where an appropriate initialization is

$$\lambda_{-2} = \lambda_{-1} = 1 \tag{6.11}$$

$$\boldsymbol{\ell}_0 = \boldsymbol{0} \tag{6.12}$$

$$\boldsymbol{w}_{-1}^D = \boldsymbol{0}, \quad \boldsymbol{w}_0^D = -\boldsymbol{e}_0^D. \tag{6.13}$$

Eqs (6.10)-(6.13) obviously yield the following algorithm:

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
\lambda_{-2} = \lambda_{-1} = 1 \\
\boldsymbol{\ell}_0 = \mathbf{0} \\
\boldsymbol{e}_0 \in \boldsymbol{\mathcal{H}} \\
\boldsymbol{w}_{-1} = \mathbf{0},\ \boldsymbol{w}_0 = -P_{\boldsymbol{D}}\, \boldsymbol{e}_0
\end{array}
\right. \\
\text{For } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\boldsymbol{v}_{n+1} = P_{\boldsymbol{D}}(\lambda_{n-1}\boldsymbol{w}_n - \boldsymbol{e}_n) \\
\boldsymbol{\ell}_{n+1} = (1-\lambda_n) P_{\boldsymbol{D}}(\boldsymbol{e}_n - \lambda_{n-2}\boldsymbol{w}_{n-1}) \\
\text{For } i = 1, \ldots, m \\
\quad \left\lfloor\; w_{i,n+1} = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i}\big((1-\varepsilon_i)(v_{i,n+1} + e_{i,n} - \ell_{i,n}) + \varepsilon_i w_{i,n}\big) + a_{i,n} \right. \\
\boldsymbol{e}_{n+1} = \boldsymbol{v}_{n+1} + \boldsymbol{e}_n - \lambda_n \boldsymbol{w}_{n+1} + \boldsymbol{\ell}_{n+1}.
\end{array}
\right.
\end{array}
\tag{6.14}
$$

By setting now $\boldsymbol{w}_0 = \boldsymbol{L}u_0$ and, for every $n \in \mathbb{N}$, $\boldsymbol{v}_{n+1} = \boldsymbol{L}u_{n+1}$ and $\boldsymbol{\ell}_{n+1} = (1-\lambda_n)\boldsymbol{L}k_{n+1}$ where $(u_n)_{n\geq 0}$ and $(k_n)_{n>0}$ are sequences of E, Algorithm (6.1) is obtained.

Since the assumptions of Proposition 5.2 hold, $\boldsymbol{x}_n = \boldsymbol{t}_n^D \rightharpoonup \boldsymbol{L}\widetilde{y}$ where $\boldsymbol{L}\widetilde{y}$ is a minimizer of $\boldsymbol{f}$ over $\boldsymbol{D}$. It can be deduced from (6.8) that $\lambda_n(\boldsymbol{w}_{n+1}^D - \boldsymbol{t}_n^D) \rightharpoonup \mathbf{0}$. In addition, from the definition of $\boldsymbol{v}_{n+1}$ and (6.7), for every $n \in \mathbb{N}^*$,

$$
\begin{aligned}
\boldsymbol{v}_{n+1} &= \lambda_{n-1}\boldsymbol{w}_n^D - \lambda_{n-2}\boldsymbol{w}_{n-1}^D + \boldsymbol{t}_n^D \\
&= \lambda_{n-1}(\boldsymbol{w}_n^D - \boldsymbol{t}_{n-1}^D) - \lambda_{n-2}(\boldsymbol{w}_{n-1}^D - \boldsymbol{t}_{n-2}^D) + \lambda_{n-1}\boldsymbol{t}_{n-1}^D - \lambda_{n-2}\boldsymbol{t}_{n-2}^D + \boldsymbol{t}_n^D.
\end{aligned}
\tag{6.15}
$$

Due to Assumption (ii) in Proposition 5.2, $\lambda_{n-1}\boldsymbol{t}_{n-1}^D - \lambda_{n-2}\boldsymbol{t}_{n-2}^D \rightharpoonup \mathbf{0}$, which allows us to conclude that $\boldsymbol{v}_{n+1} \rightharpoonup \boldsymbol{L}\widetilde{y}$. □

Similarly to Corollary 5.3, we have also:

Corollary 6.2 *Suppose that the assumptions of Proposition 5.2 hold and that* $\sum_{i=1}^m \omega_i L_i^* L_i$ *is an isomorphism. Let* $(u_n)_{n\in\mathbb{N}}$ *be a sequence generated by Algorithm* (6.1). *If the set of solutions to Problem* (5.1) *is nonempty, then* $(u_n)_{n\in\mathbb{N}}$ *converges weakly to an element of this set.*

Note that Algorithm (6.1) may appear somewhat less efficient than Algorithm (5.5) in the sense that it requires to compute two projections on $\boldsymbol{D}$ at each iteration. This disadvantage no longer exists in the unrelaxed case ($\lambda_n \equiv 1$) where Algorithm (6.1) takes a simplified form.

Proposition 6.3 *Let* $(\varepsilon_1, \dots, \varepsilon_m) \in [0,1[^m$. *Algorithm* (6.1) *with* $\lambda_n \equiv 1$ *is equivalent to the following routine:*

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
(e_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}} \\
u_0 = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z + e_{i,0}\|_i^2 \\
\text{For } i = 1, \dots, m \\
\quad \lfloor\; w_{i,0} = L_i u_0
\end{array}
\right. \\
\text{For } n = 0, 1, \dots \\
\left\lfloor
\begin{array}{l}
u_{n+1} = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - w_{i,n} + e_{i,n}\|_i^2 \\
\text{For } i = 1, \dots, m \\
\quad \left\lfloor
\begin{array}{l}
v_{i,n+1} = L_i u_{n+1} \\
w_{i,n+1} = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i} \big((1-\varepsilon_i)(v_{i,n+1} + e_{i,n}) + \varepsilon_i w_{i,n}\big) + a_{i,n} \\
e_{i,n+1} = v_{i,n+1} + e_{i,n} - w_{i,n+1}.
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{6.16}
$$

Suppose that the following assumptions hold.

(i) $(0, \dots, 0) \in \operatorname{sri} \big\{ (L_1 z - x_1, \dots, L_m z - x_m) \;\big|\; z \in E,\ x_1 \in \operatorname{dom} f_1, \dots, x_m \in \operatorname{dom} f_m \big\}$.

(ii) $(\forall i \in \{1, \dots, m\})\ \sum_{n\in\mathbb{N}} \|a_{i,n}\|_i < +\infty$.

(iii) $\sum_{i=1}^m \omega_i L_i^* L_i$ *is an isomorphism.*

If the set of solutions to Problem (5.1) *is nonempty, then* $(u_n)_{n\in\mathbb{N}}$ *converges weakly to an element of this set.*

The above algorithm when $E = \mathcal{G}$, $\varepsilon_1 = \ldots = \varepsilon_m = 0$, $\omega_1 = \ldots = \omega_m$ and $(a_{i,n})_{1\le i\le m} \equiv (0, \dots, 0)$ was derived in [45] from the Douglas-Rachford algorithm by invoking a duality argument. We have here obtained this algorithm by following a different way. In addition, the convergence result stated in Proposition 6.3 is not restricted to finite dimensional Hilbert spaces and it allows us to take into account an error term $\big((a_{i,n})_{1\le i\le m}\big)_{n\in\mathbb{N}}$ in the computation of the proximity operators of the functions $(f_i)_{1\le i\le m}$ and to include inertia parameters. It can also be noticed that Algorithm (5.5) may appear more flexible than Algorithm (6.16) since it offers the ability of using relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$.

Similarly to the derivation of Algorithm (5.11), when Assumption (iii) of Proposition 6.3 is not satisfied a variant of Algorithm (6.16) can be employed to solve Problem (5.1).

Proposition 6.4 *Let* $(\varepsilon_1,\dots,\varepsilon_m) \in [0,1[^m$ *and* $\alpha \in]0,+\infty[$. *For every* $i \in \{1,\dots,m\}$, *let* $(a_{i,n})_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}_i$. *Let* $\{u_n\}_{n\in\mathbb{N}} \subset E$, $\{\boldsymbol{v}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ $\{\boldsymbol{w}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ *and* $\{\boldsymbol{e}_n\}_{n\in\mathbb{N}} \subset \boldsymbol{\mathcal{H}}$ *be generated by the routine:*

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
(e_{i,0})_{1\le i\le m} \in \boldsymbol{\mathcal{H}}\\
u_0 = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z + e_{i,0}\|_i^2 + \alpha\|z\|^2\\
\text{For } i = 1,\dots,m\\
\quad\left\lfloor\ w_{i,0} = L_i u_0\right.
\end{array}
\right.\\
\text{For } n = 0,1,\dots\\
\left\lfloor
\begin{array}{l}
u_{n+1} = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - w_{i,n} + e_{i,n}\|_i^2 + \alpha\|z - u_n\|^2\\
\text{For } i = 1,\dots,m\\
\quad\left\lfloor
\begin{array}{l}
v_{i,n+1} = L_i u_{n+1}\\
w_{i,n+1} = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i}\big((1-\varepsilon_i)(v_{i,n+1} + e_{i,n}) + \varepsilon_i w_{i,n}\big) + a_{i,n}\\
e_{i,n+1} = v_{i,n+1} + e_{i,n} - w_{i,n+1}.
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{6.17}
$$

Suppose that Assumptions (i) *and* (ii) *of Proposition 6.3 hold. If the set of solutions to Problem* (5.1) *is nonempty, then* $(u_n)_{n\in\mathbb{N}}$ *converges weakly to an element of this set.*

Proof. Problem (5.1) can be reformulated as

$$
\underset{y\in E}{\text{minimize}} \sum_{i=1}^{m+1} f_i(L_i y) \tag{6.18}
$$

where $f_{m+1} = 0$ and $L_{m+1} = \operatorname{Id}$. Hence, for every $\omega_{m+1} \in]0,+\infty[$, $\sum_{i=1}^{m+1} \omega_i L_i^* L_i = \sum_{i=1}^m \omega_i L_i^* L_i + \omega_{m+1} \operatorname{Id}$ is an isomorphism. By applying a version of Algorithm (6.16) to the above problem with $m+1$ potentials, we get:

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
(e_{i,0})_{1\le i\le m+1} \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_m \times \mathcal{G}\\
u_0 = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z + e_{i,0}\|_i^2 + \omega_{m+1}\|z + e_{m+1,0}\|^2\\
\text{For } i = 1,\dots,m+1\\
\quad\left\lfloor\ w_{i,0} = L_i u_0\right.
\end{array}
\right.\\
\text{For } n = 0,1,\dots\\
\left\lfloor
\begin{array}{l}
u_{n+1} = \operatorname{argmin}_{z\in E} \sum_{i=1}^m \omega_i \|L_i z - w_{i,n} + e_{i,n}\|_i^2 + \omega_{m+1}\|z - w_{m+1,n} + e_{m+1,n}\|^2\\
\text{For } i = 1,\dots,m+1\\
\quad\left\lfloor
\begin{array}{l}
v_{i,n+1} = L_i u_{n+1}\\
w_{i,n+1} = \operatorname{prox}_{\frac{1-\varepsilon_i}{\omega_i} f_i}\big((1-\varepsilon_i)(v_{i,n+1} + e_{i,n}) + \varepsilon_i w_{i,n}\big) + a_{i,n}\\
e_{i,n+1} = v_{i,n+1} + e_{i,n} - w_{i,n+1}.
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{6.19}
$$

Then, by setting $e_{m+1,0} = 0$, $\omega_{m+1} = \alpha$, $\varepsilon_{m+1} = 0$, $a_{m+1,n} \equiv 0$, and noticing that $(\forall n \in \mathbb{N})$ $e_{m+1,n+1} = 0$, $v_{m+1,n+1} = u_{n+1}$ and $w_{m+1,n} = u_n$, Algorithm (6.17) is obtained.

In addition, Assumption (i) in Proposition 6.3 is equivalent to assume that

$$
\boldsymbol{V} = \bigcup_{\lambda>0} \big\{\lambda(L_1 z - x_1,\dots,L_m z - x_m) \bigm| z \in E,\ x_1 \in \operatorname{dom} f_1,\dots,x_m \in \operatorname{dom} f_m\big\}
$$

is a closed vector subspace of $\boldsymbol{\mathcal{H}}$. $\boldsymbol{V}\times\mathcal{G}$ is therefore a closed vector subspace of $\boldsymbol{\mathcal{H}}\times\mathcal{G}$. But, since $\operatorname{dom} f_{m+1}=\mathcal{G}$, we have

$$\boldsymbol{V}\times\mathcal{G}=\bigcup_{\lambda>0}\{\lambda(L_1z-x_1,\ldots,L_mz-x_m,z-x_{m+1})\mid$$
$$z\in E,\ x_1\in\operatorname{dom} f_1,\ldots,\ x_m\in\operatorname{dom} f_m, x_{m+1}\in\operatorname{dom} f_{m+1}\}.$$

This shows that

$$\mathbf{0}\in\operatorname{sri}\big\{(L_1z-x_1,\ldots,L_mz-x_m,z-x_{m+1})\;\big|\;z\in E,\ x_1\in\operatorname{dom} f_1,\ldots,x_{m+1}\in\operatorname{dom} f_{m+1}\big\}. \tag{6.20}$$

Since this condition is satisfied and Assumptions (ii) and (iii) hold, the convergence of $(u_n)_{n\in\mathbb{N}}$ follows from Proposition 6.3 applied to Problem (6.18). ◻

Remark 6.5

(i) In the case when $\varepsilon_1=\ldots=\varepsilon_m=0$, Algorithm (6.16) can be derived in a more direct way from Spingarn's method of partial inverses [46, 47], which is recalled below:

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}(\boldsymbol{s}_0,\boldsymbol{q}_0)\in\boldsymbol{D}\times\boldsymbol{D}^\perp\end{array}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
(\boldsymbol{w}'_{n+1},\boldsymbol{e}'_{n+1})\text{ such that }\boldsymbol{w}'_{n+1}+\boldsymbol{e}'_{n+1}=\boldsymbol{s}_n+\boldsymbol{q}_n\text{ and }\boldsymbol{e}'_{n+1}\in\partial\boldsymbol{f}(\boldsymbol{w}'_{n+1})\\
\boldsymbol{s}_{n+1}=P_{\boldsymbol{D}}\boldsymbol{w}'_{n+1}\\
\boldsymbol{q}_{n+1}=P_{\boldsymbol{D}^\perp}\boldsymbol{e}'_{n+1}
\end{array}\right.
\end{array} \tag{6.21}$$

Indeed at each iteration $n\in\mathbb{N}$, $\boldsymbol{w}'_{n+1}$ and $\boldsymbol{e}'_{n+1}$ are then computed as

$$\boldsymbol{w}'_{n+1}=\operatorname{prox}_{\boldsymbol{f}}(\boldsymbol{s}_n+\boldsymbol{q}_n) \tag{6.22}$$
$$\boldsymbol{e}'_{n+1}=\boldsymbol{s}_n+\boldsymbol{q}_n-\boldsymbol{w}'_{n+1}. \tag{6.23}$$

By setting $\boldsymbol{w}'_0=\boldsymbol{s}_0$ and $\boldsymbol{e}'_0=\boldsymbol{q}_0-\boldsymbol{s}_0$, and by defining, for every $n\in\mathbb{N}$,

$$\boldsymbol{v}'_{n+1}=\boldsymbol{s}_n-P_{\boldsymbol{D}}\boldsymbol{e}'_n=P_{\boldsymbol{D}}(\boldsymbol{w}'_n-\boldsymbol{e}'_n) \tag{6.24}$$

we have

$$\boldsymbol{s}_n+\boldsymbol{q}_n=\boldsymbol{v}'_{n+1}+P_{\boldsymbol{D}}\boldsymbol{e}'_n+P_{\boldsymbol{D}^\perp}\boldsymbol{e}'_n=\boldsymbol{v}'_{n+1}+\boldsymbol{e}'_n. \tag{6.25}$$

Altogether, (6.22)-(6.25) yield the following algorithm:

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}\boldsymbol{e}'_0\in\boldsymbol{\mathcal{H}}\\ \boldsymbol{w}'_0=-P_{\boldsymbol{D}}\boldsymbol{e}'_0\end{array}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
\boldsymbol{v}'_{n+1}=P_{\boldsymbol{D}}(\boldsymbol{w}'_n-\boldsymbol{e}'_n)\\
\boldsymbol{w}'_{n+1}=\operatorname{prox}_{\boldsymbol{f}}(\boldsymbol{v}'_{n+1}+\boldsymbol{e}'_n)\\
\boldsymbol{e}'_{n+1}=\boldsymbol{v}'_{n+1}+\boldsymbol{e}'_n-\boldsymbol{w}'_{n+1}.
\end{array}\right.
\end{array} \tag{6.26}$$

By setting $(\boldsymbol{w}'_n)_{n\in\mathbb{N}}=(\boldsymbol{w}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{e}'_n)_{n\in\mathbb{N}}=(\boldsymbol{e}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{v}'_n)_{n>0}=(\boldsymbol{v}_n)_{n>0}$, the above algorithm leads to Algorithm (6.16) when $\varepsilon_1=\ldots=\varepsilon_m=0$, in the absence of error terms.

(ii) When $\varepsilon_1 = \ldots = \varepsilon_m = \alpha/(1+\alpha)$ and $\boldsymbol{a}_n \equiv 0$, the iterations of Algorithm (6.17) can be re-expressed as

$$(\forall n \in \mathbb{N}) \qquad u_{n+1} = \mathrm{argmin}_{z \in E} \, |||\boldsymbol{L}z - \boldsymbol{w}_n + \boldsymbol{e}_n|||^2 + \alpha \|z - u_n\|^2 \tag{6.27}$$

$$\boldsymbol{w}_{n+1} = \mathrm{prox}_{\frac{\boldsymbol{f}}{1+\alpha}} \Big(\frac{1}{1+\alpha}(\boldsymbol{L}u_{n+1} + \boldsymbol{e}_n) + \frac{\alpha}{1+\alpha}\boldsymbol{w}_n \Big) \tag{6.28}$$

$$\boldsymbol{e}_{n+1} = \boldsymbol{L}u_{n+1} + \boldsymbol{e}_n - \boldsymbol{w}_{n+1} \tag{6.29}$$

which is equivalent to

$$(\forall n \in \mathbb{N}) \quad u_{n+1} = \mathrm{argmin}_{z \in \mathcal{G}} \, \iota_E(z) + |||\boldsymbol{L}z - \boldsymbol{w}_n + \boldsymbol{e}_n|||^2 + \alpha \|z - u_n\|^2 \tag{6.30}$$

$$\begin{aligned} \boldsymbol{w}_{n+1} &= \mathrm{argmin}_{\boldsymbol{x} \in \boldsymbol{\mathcal{H}}} \, \frac{1}{1+\alpha}\boldsymbol{f}(\boldsymbol{x}) + \frac{1}{2}\, |||\boldsymbol{x} - \frac{1}{1+\alpha}(\boldsymbol{L}u_{n+1} + \boldsymbol{e}_n) - \frac{\alpha}{1+\alpha}\boldsymbol{w}_n|||^2 \\ &= \mathrm{argmin}_{\boldsymbol{x} \in \boldsymbol{\mathcal{H}}} \, \boldsymbol{f}(\boldsymbol{x}) + \frac{1}{2}|||\boldsymbol{x} - \boldsymbol{L}u_{n+1} - \boldsymbol{e}_n|||^2 + \frac{\alpha}{2}|||\boldsymbol{x} - \boldsymbol{w}_n|||^2 \end{aligned} \tag{6.31}$$

$$\boldsymbol{e}_{n+1} = \boldsymbol{L}u_{n+1} + \boldsymbol{e}_n - \boldsymbol{w}_{n+1}. \tag{6.32}$$

By setting $\gamma = \alpha^{-1/2}$ and $(\forall n > 0)$ $\boldsymbol{e}'_n = \gamma \boldsymbol{e}_n$, these iterations read also

$$(\forall n \in \mathbb{N}) \quad u_{n+1} = \mathrm{argmin}_{z \in \mathcal{G}} \, \iota_E(z) + \langle\langle \boldsymbol{e}'_n \mid \boldsymbol{L}z \rangle\rangle + \frac{\gamma}{2}|||\boldsymbol{L}z - \boldsymbol{w}_n|||^2 + \frac{1}{2\gamma}\|z - u_n\|^2 \tag{6.33}$$

$$\boldsymbol{w}_{n+1} = \mathrm{argmin}_{\boldsymbol{x} \in \boldsymbol{\mathcal{H}}} \, \gamma \boldsymbol{f}(\boldsymbol{x}) - \langle\langle \boldsymbol{e}'_n \mid \boldsymbol{x} \rangle\rangle + \frac{\gamma}{2}|||\boldsymbol{x} - \boldsymbol{L}u_{n+1}|||^2 + \frac{1}{2\gamma}|||\boldsymbol{x} - \boldsymbol{w}_n|||^2 \tag{6.34}$$

$$\boldsymbol{e}'_{n+1} = \boldsymbol{e}'_n + \gamma(\boldsymbol{L}u_{n+1} - \boldsymbol{w}_{n+1}). \tag{6.35}$$

Therefore, this algorithm takes the form of the one proposed in [5] when applied to the optimization problem

$$\underset{y \in \mathcal{G}}{\mathrm{minimize}} \;\; \iota_E(y) + \gamma \boldsymbol{f}(\boldsymbol{L}y). \tag{6.36}$$

In other words, when $\varepsilon_1 = \ldots = \varepsilon_m = \alpha/(1+\alpha)$ and $\boldsymbol{a}_n \equiv 0$, Algorithm (6.17) can be viewed as a parallelized version of the algorithm proposed in [5].

7 Numerical experiments

In order to evaluate the performance of the proposed algorithm, we present numerical experiments relative to an image restoration problem. The considered degradation model is $z = A\overline{y} + \epsilon$ where $\overline{y} \in \mathbb{R}^N$ denotes the image to be recovered, $z \in \mathbb{R}^N$ are the observed data, $A \in \mathbb{R}^{N \times N}$ corresponds to a circulant-block circulant matrix associated with a blur operator, and $\epsilon \in \mathbb{R}^N$ is a realization of a random vector with independent and identically distributed components following a uniform distribution. The restoration process is based on a variational approach aiming at finding an image $\widehat{y} \in \mathbb{R}^N$ as close as possible to $\overline{y} \in \mathbb{R}^N$ from the observation z, and prior informations such as the dynamic range of the pixel values and the sparsity of $\overline{y}$ in a wavelet-frame representation. The considered optimization problem is :

$$\text{find} \quad \widehat{y} \in \underset{y \in \mathbb{R}^N}{\mathrm{Argmin}} \, \|z - Ay\|_{\ell_p}^p + \vartheta \|Fy\|_{\ell_1} + \iota_C(y) \tag{7.1}$$

with $p \geq 1$ and $\vartheta > 0$. $F \in \mathbb{R}^{K \times N}$ denotes the matrix associated with the frame analysis operator. The frame is supposed to be tight, i.e. $F^*F = \nu_F \,\mathrm{Id}$, for some $\nu_F > 0$. $C = [0,255]^N$ models a dynamic range constraint. The notation $\|\cdot\|_{\ell_p}$ stands for the classical ℓ_p norm:

$$(\forall y = (\eta_i)_{1\leq i\leq N} \in \mathbb{R}^N) \qquad \|y\|_{\ell_p} = \Big(\sum_{i=1}^{N} |\eta_i|^p\Big)^{1/p}. \tag{7.2}$$

Problem (7.1) is a particular case of Problem (5.1) when $m = 3$, $E = \mathbb{R}^N$, $f_1 = \|z - \cdot\|_{\ell_p}^p$, $f_2 = \|\cdot\|_{\ell_1}$, $f_3 = \iota_C$, $L_1 = A$, $L_2 = F$, and $L_3 = \mathrm{Id}$.

In the following, this approach is used to restore a satellite image of size $N = 256\times256$ corrupted by a Gaussian blur with a large kernel size and a noise uniformly distributed on $[-30, 55]$. The latter values are not assumed to be known in the recovery process. In order to deal efficiently with the uniformly distributed noise, the data fidelity measure f_1 is chosen to be an ℓ_3-norm ($p = 3$). F corresponds to a tight frame version of the dual-tree transform (DTT) proposed in [14] ($\nu_F = 2$) using symmlets of length 6 over 2 resolution levels.

Algorithm (5.5) is then employed to solve (7.1). The proximity operator of f_1 here takes a closed form expression [13], whereas the proximity operator of f_2 reduces to a soft-thresholding rule (see [20] and references therein). In order to efficiently compute c_n at iteration n of (5.5), fast Fourier diagonalization techniques have been employed. It can be noticed that the qualification condition (i) in Proposition 5.2 is satisfied since $\mathrm{dom}\, f_1 = \mathrm{dom}\, f_2 = \mathbb{R}^N$ and $\mathrm{int}\, C = \,]0, 255[^N \neq \varnothing$ (see Remark 5.5(i)). The convergence to an optimal solution to Problem 7.1 is then guaranteed by Corollary 5.3.

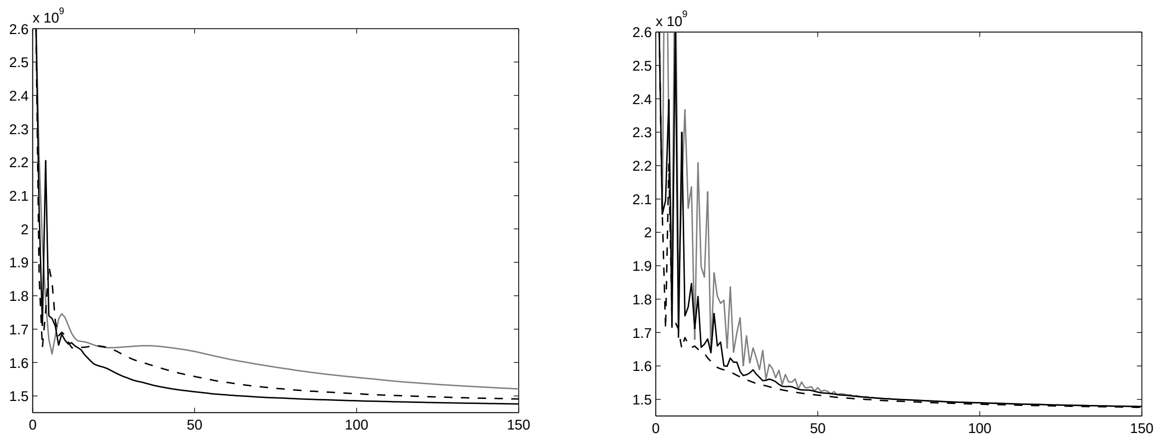
The original, degraded, and restored images are presented in Figure 1. The value of ϑ has been selected so as to maximize the signal-to-noise-ratio (SNR). The SNR between an image y and the original image $\overline{y}$ is defined as $20\log_{10}(\|\overline{y}\|_{\ell_2}/\|y - \overline{y}\|_{\ell_2})$. In our experiments, we also compare the images in terms of structural similarity (SSIM) [50]. SSIM can be seen as a percentage of similarity, the value 1 being achieved for two identical images.

Note that for such an application, the Douglas-Rachford algorithm would not be a proper choice since there is no closed form expression for the proximity operator of $f_1 \circ A$. Regarding PPXA [21, Algorithm 3.1], it would be possible to split the data fidelity term as proposed in [43], but this solution would not be as efficient as the proposed one due to the large kernel size of the blur.

The effects of the parameters $(\lambda_n)_{n\in\mathbb{N}}$ and $(\varepsilon_i)_{1\leq i\leq m}$ are evaluated in Figures 2(a) and (b) where the criterion variations $(\|z - Ay_n\|_{\ell_p}^p + \vartheta\|Fy_n\|_{\ell_1})_n$ are plotted as a function of the iteration number n. On the one hand, it results from our experiments that for a fixed relaxation parameter $\lambda_n \equiv \lambda$, a faster convergence is achieved by choosing λ close to 2. Recall that, according to Proposition 6.3, there exist close connections between the proposed algorithm and SDMM when $\lambda = 1$. On the other hand, an identical value $\varepsilon = 0.4$ of the inertial parameters $(\varepsilon_i)_{1\leq i\leq m}$ also appears to be beneficial to the convergence profile. In particular, the introduction of inertial parameters allows us to reduce the oscillations in the variations of the criterion.

Original | Degraded SNR = 8.56 dB SSIM = 0.43 | Restored SNR = 12.4 dB SSIM = 0.71

Figure 1: Image restoration result.



(a) (b)

Figure 2: Impact of the algorithm parameters on the convergence rate. (a) Relaxation parameter when $\epsilon = 0.4$ and $\lambda_n \equiv \lambda$ with $\lambda = 0.5$ (gray solid line), $\lambda = 1$ (black dashed line), and $\lambda = 1.9$ (black solid line). (b) Inertial parameters when $\lambda = 1.9$ and $\varepsilon_i \equiv \varepsilon$ with $\varepsilon = 0$ (gray solid line), $\varepsilon = 0.4$ (black dashed line), and $\varepsilon = 0.8$ (black solid line).

Appendix

Proposition 7.1 *Let $\boldsymbol{g}$ be the function defined as*

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) + \frac{\kappa}{2}|||\boldsymbol{x}|||^2 \tag{7.3}$$

where $\boldsymbol{h} \in \Gamma_0(\boldsymbol{\mathcal{H}})$ has a β-Lipschitz continuous gradient and $(\beta, \kappa) \in]0, +\infty[^2$. Then $\mathrm{rprox}_{\boldsymbol{g}}$ *is strictly contractive with constant* $(1+\kappa)^{-1}\big((1-\kappa)^2 + 4\kappa(1+\beta(1+\kappa))^{-1}\big)^{1/2}$.

Proof. From (7.3) and standard properties of the proximity operator [22, 23], we get

$$\begin{aligned}(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad \operatorname{prox}_{\boldsymbol{g}} \boldsymbol{x} &= \operatorname{prox}_{\frac{\boldsymbol{h}}{1+\kappa}} \Big(\frac{\boldsymbol{x}}{1+\kappa}\Big) \\ &= \frac{1}{1+\kappa}\Big(\boldsymbol{x} - \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*} \boldsymbol{x}\Big) \\ \Leftrightarrow \quad \operatorname{rprox}_{\boldsymbol{g}} \boldsymbol{x} &= \frac{1}{1+\kappa}\Big((1-\kappa)\boldsymbol{x} - 2\operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*} \boldsymbol{x}\Big). \end{aligned} \tag{7.4}$$

Then, we have

$$\begin{aligned}&(\forall (\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}}^2) \quad |||\operatorname{rprox}_{\boldsymbol{g}} \boldsymbol{x} - \operatorname{rprox}_{\boldsymbol{g}} \boldsymbol{y}|||^2 \\ &\quad = \frac{1}{(1+\kappa)^2}\Big((1-\kappa)^2|||\boldsymbol{x}-\boldsymbol{y}|||^2 - 4(1-\kappa)\langle\langle \boldsymbol{x}-\boldsymbol{y} \mid \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{x} - \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{y}\rangle\rangle \\ &\qquad + 4\,|||\operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{x} - \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{y}|||^2\Big) \\ &\quad \le \frac{1}{(1+\kappa)^2}\Big((1-\kappa)^2|||\boldsymbol{x}-\boldsymbol{y}|||^2 + 4\kappa\langle\langle \boldsymbol{x}-\boldsymbol{y} \mid \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{x} - \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{y}\rangle\rangle\Big)\end{aligned} \tag{7.5}$$

where the fact that the proximity operator is firmly nonexpansive has been used in the last inequality. On the other hand, according to [6, Theorem 2.1], we know that $\boldsymbol{h}$ has a $1/\beta$-Lipschitz continuous gradient if and only if $\boldsymbol{h}^*$ is strongly convex with modulus β, that is

$$\boldsymbol{h}^* = \boldsymbol{\varphi} + \frac{\beta}{2}|||\cdot|||^2. \tag{7.6}$$

where $\boldsymbol{\varphi} \in \Gamma_0(\boldsymbol{\mathcal{H}})$. Thus,

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \qquad \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{x} = \operatorname{prox}_{\frac{(1+\kappa)\boldsymbol{\varphi}}{1+\beta(1+\kappa)}}\Big(\frac{\boldsymbol{x}}{1+\beta(1+\kappa)}\Big). \tag{7.7}$$

By invoking now the Cauchy-Schwarz inequality and the nonexpansivity of $\operatorname{prox}_{\frac{(1+\kappa)\boldsymbol{\varphi}}{1+\beta(1+\kappa)}}$, we deduce that

$$\begin{aligned}&\langle\langle \boldsymbol{x}-\boldsymbol{y} \mid \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{x} - \operatorname{prox}_{(1+\kappa)\boldsymbol{h}^*}\boldsymbol{y}\rangle\rangle \\ &\le |||\boldsymbol{x}-\boldsymbol{y}|||\;\Big|\Big|\Big|\operatorname{prox}_{\frac{(1+\kappa)\boldsymbol{\varphi}}{1+\beta(1+\kappa)}}\Big(\frac{\boldsymbol{x}}{1+\beta(1+\kappa)}\Big) - \operatorname{prox}_{\frac{(1+\kappa)\boldsymbol{\varphi}}{1+\beta(1+\kappa)}}\Big(\frac{\boldsymbol{y}}{1+\beta(1+\kappa)}\Big)\Big|\Big|\Big| \\ &\le \frac{1}{1+\beta(1+\kappa)}|||\boldsymbol{x}-\boldsymbol{y}|||^2.\end{aligned} \tag{7.8}$$

By combining (7.5) and (7.8), we finally obtain

$$\begin{aligned}(\forall (\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}}^2) \quad &|||\operatorname{rprox}_{\boldsymbol{g}} \boldsymbol{x} - \operatorname{rprox}_{\boldsymbol{g}} \boldsymbol{y}|||^2 \\ &\le \frac{1}{(1+\kappa)^2}\Big((1-\kappa)^2 + \frac{4\kappa}{1+\beta(1+\kappa)}\Big)|||\boldsymbol{x}-\boldsymbol{y}|||^2.\end{aligned} \tag{7.9}$$

Since

$$\frac{1}{(1+\kappa)^2}\Big((1-\kappa)^2 + \frac{4\kappa}{1+\beta(1+\kappa)}\Big) < \frac{1}{(1+\kappa)^2}\big((1-\kappa)^2 + 4\kappa\big) = 1 \tag{7.10}$$

$\operatorname{rprox}_{\boldsymbol{g}}$ is thus strictly contractive. □

References

[1] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.*, 19(9):2345–2356, Sep. 2010. http://arxiv.org/abs/0910.4887.

[2] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.*, 20(3):681–695, March 2011.

[3] F. Alvarez. Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. *SIAM J. Optim.*, 14:773–782, 2003.

[4] F. Alvarez and H. Attouch. Convergence and asymptotic stabilization for some damped hyperbolic equations with non-isolated equilibria. *ESAIM Control Optim. Calc. Var.*, 6:539–552, 2001.

[5] H. Attouch and M. Soueycatt. Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces – applications to games, PDEs and control. *Pacific J. Optim.*, 5:17–37, 2009.

[6] H. H. Bauschke and P. L. Combettes. The Baillon-Haddad theorem revisited. *Journal of Convex Analysis*, 17(4), December 2010.

[7] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer-Verlag, New York, 2011.

[8] H.H. Bauschke. A note on the paper by Eckstein and Svaiter on "General projective splitting methods for sums of maximal monotone operators". *SIAM J. Control Optim.*, 48:2513–2515, 2009.

[9] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods.* Athena Scientific, Belmont, MA, 1997.

[10] L. M. Briceño-Arias and P. L. Combettes. Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery. *Numer. Math. Theory Methods Appl.*, 2:485–508, 2009.

[11] F. E. Browder. Convergence theorems for sequences of nonlinear operators in Banach spaces. *Mathematische Zeitschrift*, 100:201–225, 1967.

[12] Y. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms and Applications.* Oxford University Press, New York, 1997.

[13] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A variational formulation for frame-based inverse problems. *Inverse Problems*, 23:1495–1518, 2007.

[14] C. Chaux, L. Duval, and J.-C. Pesquet. Image analysis using a dual-tree M-band wavelet transform. *IEEE Trans. Image Process.*

[15] C. Chaux, J.-C. Pesquet, and N. Pustelnik. Nested iterative algorithms for convex constrained image recovery problems. *SIAM J. Imaging Sci.*, 2:730–762, 2009.

[16] P. L. Combettes. Quasi-Fejérian analysis of some optimization algorithms. In D. Butnariu, Y. Censor, and S. Reich, editors, *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, pages 115–152. Elsevier, New York, 2001.

[17] P. L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53:475–504, 2004.

[18] P. L. Combettes. Iterative construction of the resolvent of a sum of maximal monotone operators. *J. Convex Anal.*, 16:727–748, 2009.

[19] P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Selected Topics Signal Process.*, 1:564–574, 2007.

[20] P. L. Combettes and J.-C. Pesquet. Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM J. Optim.*, 18:1351–1376, 2007.

[21] P. L. Combettes and J.-C. Pesquet. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, 24, 2008. 27 p.,art. 065014.

[22] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer-Verlag, New York, 2010.

[23] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, 4:1168–1200, 2005.

[24] J. Eckstein. Parallel alternating direction multiplier decomposition of convex programs. *J. Optim. Theory Appl.*, 80:39–62, 1994.

[25] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 55:293–318, 1992.

[26] J. Eckstein and B. F. Svaiter. A family of projective splitting methods for the sum of two maximal monotone operators. *Math. Programming*, 111:173–199, 2008.

[27] J. Eckstein and B. F. Svaiter. General projective splitting methods for sums of maximal monotone operators. *SIAM J. Control Optim.*, 48:787–811, 2009.

[28] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. SIAM, Philadelphia, PA, 1999.

[29] E. Esser. Applications of Lagrangian-based alternating direction methods and connections to split Bregman. ftp://ftp.math.ucla.edu/pub/camreport/cam09-31.pdf, 2009.

[30] M. Fortin and R. Glowinski, editors. *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*. North-Holland, Amsterdam, 1983.

[31] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite elements approximations. *Comput. Math. Appl.*, 2:17–40, 1976.

[32] R. Glowinski and P. Le Tallec, editors. *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM, Philadelphia, 1989.

[33] T. Goldstein and S. Osher. The split Bregman method for $L1$-regularized problems. *SIAM J. Imaging Sci.*, 2:323–343, 2009.

[34] S. P. Han and G. Lou. A parallel algorithm for a class of convex programs. *SIAM J. Control Optim*, 26:345–355, 1988.

[35] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16:964–979, 1979.

[36] P.-E. Maingé. Inertial iterative process for fixed points of certain quasi-nonexpansive mappings. *Set-Valued Anal.*, 15(1):67–79, 2007.

[37] P.-E. Maingé. Convergence theorems for inertial KM-type algorithms. *J. Comput. Appl. Math.*, 219(1):223–236, 2008.

[38] P.-E. Maingé. Regularized and inertial algorithms for common fixed points of nonlinear operators. *J. Math. Anal. Appl.*, 344(2):876–887, 2008.

[39] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[40] A. Moudafi. A hybrid inertial projection-proximal method for variational inequalities. *Journal of Inequalities in Pure and Applied Mathematics*, 5(3):1–5, 2004. Art. 63.

[41] A. Moudafi and E. Elisabeth. An approximate inertial proximal method using enlargement of a maximal monotone operator. *Int. J. Pure Appl. Math.*, 5(3):283–299, 2003.

[42] A. Moudafi and M. Oliny. Convergence of a splitting inertial proximal method for monotone operators. *Journal of Computational and Applied Mathematics*, 155:447–454, 2003.

[43] N. Pustelnik, C. Chaux, and J.-C. Pesquet. Parallel proximal algorithm for image restoration using hybrid regularization. to appear in IEEE Trans. Image Process., http://arxiv.org/abs/0911.1536, 2011.

[44] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, 14:877–898, 1976.

[45] S. Setzer, G. Steidl, and T. Teuber. Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image Represent.*, 21:193–199, 2010.

[46] J. E. Spingarn. Partial inverse of a monotone operator. *Appl. Math. Optim.*, 10:247–265, 1983.

[47] J. E. Spingarn. Applications of the method of partial inverses to convex programming: decomposition. *Math. Programming*, 32:199–223, 1985.

[48] G. Steidl and T. Teuber. Removing multiplicative noise by Douglas-Rachford splitting methods. *J. Math. Imaging Vision*, 36:168–184, 2010.

[49] B. F. Svaiter. On weak convergence of the Douglas-Rachford method. *SIAM J. Optim.*, 49(1):280–287, 2011.

[50] Z. Wang and A. C. Bovik. Mean squared error: love it or leave it? *IEEE Signal Process. Mag.*, 26(1):98–117, Jan. 2009.

[51] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for l1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008.

[52] X. Zhang, M. Burger, X. Bresson, and S. Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM J. Imaging Sci.*, 3:253–276, 2010. ftp://ftp.math.ucla.edu/pub/camreport/cam09-03.pdf.

[53] C. Zălinescu. *Convex Analysis in General Vector Spaces.* World Scientific, River Edge, NJ, 2002.